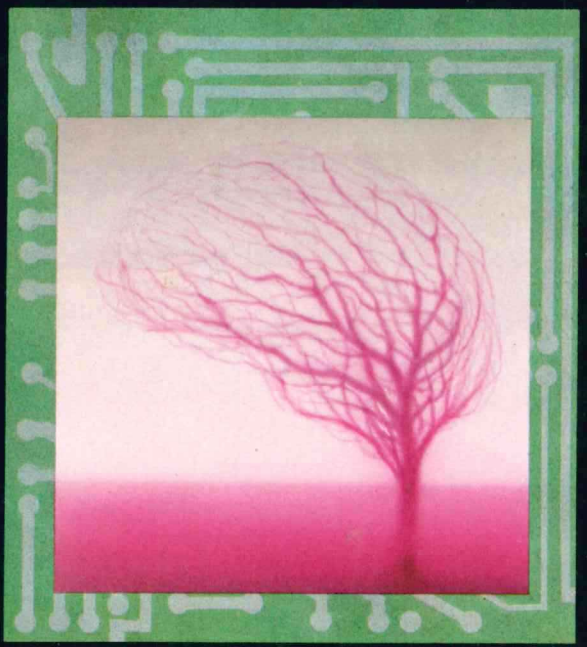
JOSEPH WEIZENBAUM

IL POTERE DEL COMPUTER E LA RAGIONE UMANA

I LIMITI DELL'INTELLIGENZA ARTIFICIALE



EDIZIONI GRUPPO ABELE

ALTRISAGGI

1. Johan Galtung
   Ci sono alternative!

2. Rodolfo Venditti
   Le ragioni dell'obiezione di coscienza

3. Francesco Alberoni - Franco Ferrarotti - Claudio Calvaruso
   I giovani verso il duemila

4. George Mc Robie
   Piccolo è possibile

Joseph Weizenbaum è nato nel 1923 a Berlino da genitori ebrei. Dopo l'ascesa al potere di Hitler emigrò negli Stati Uniti dove frèquentò le scuole pubbliche e completò i suoi studi matematici alla Wayne University di Detroit, Michigan. Qui collaborò come consulente scientifico all'ideazione e costruzione di uno dei primi computer nonché all'elaborazione di un linguaggio per la sua programmazione. Si è imposto all'attenzione dei suoi colleghi come autore di SLIP, un linguaggio di *list-processing,* e di ELIZA, un sistema di elaborazione del linguaggio naturale. Dal 1963 insegna Scienza del computer al Massachusetts Institute of Technology (MIT). È membro dell'Associazione americana per il progresso della scienza e uno dei fondatori del Computers for Social Responsability (Professionisti dell'informatica per una responsabilità sociale) negli Stati Uniti e nella Germania Federale.

Titolo originale:
Computer Power and Human Reason
From Judgment to Calculation

© 1976 W. H. Freeman and Company

Traduzione e cura redazionale di
Federico Tibone

Copertina di
Walter Oglino

Illustrazione in copertina di
Franco Bolzoni

I edizione: marzo 1987

© 1987
Edizioni Gruppo Abele

Via Giolitti, 21 - 10123 Torino - Tel. (011) 8395444
ISBN 88-7670-082-X

INTRODUZIONE ALL'EDIZIONE ITALIANA

di Franco La Cecla*

Mentre leggevo in treno le bozze di questo libro, un signore, dotato di molto buon senso, mi ha domandato cosa fosse. Alla risposta che si trattava di un' introduzione critica alla società dei computer, il signore ha sbirciato oltre i fogli e ha osservato che il testo era stato battuto al computer. Gli è sembrato di avermi fatto cadere in aperta contraddizione e per il resto del viaggio ha mantenuto un atteggiamento di paterna comprensione.

Il signore è un buon esempio dell'accoglienza che la società italiana ha nei confronti della rivoluzione informatica: nessun entusiasmo, solo un apparente realismo — «I computer ci sono, si vanno sempre più diffondendo, fanno sempre più cose utili, perché criticarli?».

Posta in questi termini, ogni questione riguardante il cambiamento che l'informatica sta inducendo nella società appare prossimo al dibattito tra vegetariani e carnivori; si può rispettare chi non mangia carne, ma perché dover accettare la sua personalissima visione critica?

Sembra che ci si trovi dinanzi ad un dibattito di natura ideologica — computer sì, computer no — come questione di fede o meno.

La povertà del dibattito è dovuta in buona parte al solito ritardo italico. Da noi sono state tradotte gran parte delle opere riguardanti la manualistica da computer, ma ben poche aventi a che fare con la rivoluzione informatica ed i suoi effetti scientifici, sociali, antropologici.

Questo libro di Joseph Weizenbaum è uscito una prima volta nel 1975 a San Francisco. Ragioni di politica editoriale ed industriale hanno sbarrato il passo alla sua traduzione per oltre dieci anni. Se si pensa che anche i suoi detrattori statunitensi sono convinti che l'opera di Weizenbaum è comunque una pietra miliare di confronto, si deduce che in Italia qualunque critica alla rivoluzione informatica viene vissuta con il complesso di Cenerentola. Sempre preoccupati di stare al passo con i tempi, sembra non ci sia tempo, anche in ambiti filosofici, scientifici ed accademici per aprire un dibattito.

Il libro, che qui compare nella traduzione di un' edizione riveduta ed ampliata dall'autore nel 1984 per la Penguin, è in realtà molto sostanzioso. Jo-

* *Franco La Cecla* è ricercatore all'Università di Bologna presso l'Istituto di Comunicazione (DAMS) della Facoltà di Lettere e pari dell'Elmwood Institute di Fritjof Capra a Berkeley.

seph Weizenbaum non è un «polemista», un sociologo alla Alberoni o alla Bocca, non è un filosofo o un umanista. Egli è uno dei fondatori della Artificial Intelligence negli Stati Uniti. A tutt'oggi insegna «Computer Science» nel prestigioso Massachusetts Institute of Technology di Boston. Ha scritto *Il potere del computer e la ragione umana* nel corso di due anni di ricerca tra Stanford, Harvard e MIT, finanziato dalla National Science Foundation. Non è dunque un ricercatore «selvaggio». Anzi è proprio il suo carattere di «interno» alla Computer Science a fare imbestialire i suoi colleghi. Chi voglia divertirsi di fronte ai toni accesissimi del dibattito tra Weizenbaum ed i suoi colleghi può leggerne il resoconto nel suo *Once more: the Computer Revolution*, in M. Dertounzos-J. Moses (a cura di), *The Computer Age. A twenty Years Rewiew*, MIT, Boston 1983, pp. 439-458.

Fatto sta che il libro che il lettore italiano si trova tra le mani non solo non è retrodotato per il contesto italiano, ma è ancora fresco ed «irritante» per quello di origine.

Weizenbaum è ritenuto una voce autorevole a tal punto che il prestigioso «New York Review of Books» gli lascia ampio spazio per commentare la nuova letteratura nel campo della Intelligenza Artificiale. E Weizenbaum lo fa con lo stile che potrete qui apprezzare: appassionato, contenuto, pungente, ma allo stèsso tempo rigorosamente disciplinare.

Il potere del computer e la ragione umana è anzitutto un' introduzione alla natura dell'informatica e dell' Intelligenza Artificiale. I non iniziati vi possono trovare un filo per accedere ad un sapere che anche in Italia rimane frammentato, sparso, dato per scontato. La povertà del dibattito italiano è tale anche perché dei computer tutto sommato si sa ben poco e si va avanti tra ovvietà. Ciò che Weizenbaum ha di anomalo, rispetto alla «disciplina» è la sua vastissima cultura, che spazia dalla linguistica alla psicologia cognitiva, dalla storia delle innovazioni tecnologiche alla storia della politica della tecnologia negli Stati Uniti. Egli rappresenta un mondo, insieme forse a personaggi come Annah Arendt, Noam Chomsky, Lewis Mumford, che non rinuncia a giudicare un presente in cui è coinvolto, di cui sta facendo parte attiva, all'interno di una disciplina e di una pratica politica. Un mondo che sente il compito di confrontare con la «cultura» e con la «politica» ogni innovazione sociale e tecnologica. Probabilmente un corrispettivo del genere è difficile trovarlo in Italia. Da quali parti, all'interno del mondo accademico, o tra gli «esperti» del settore si alzano discorsi preoccupati dell'impatto tra tecnologia e società? Forse il primo ed unico caso rimane quello dei fisici rispetto al nucleare. Ma in genere il *Question Technology* rimane un appello disatteso, o ampiamente censurato da una martellante pubblicità che precede e prepara il campo ad ogni innovazione. Sarebbe interessante ricostruire la storia delle campagne pubblicitarie ed editoriali per preparare il campo commerciale ed ideale alla introduzione dell' informatica in Italia. Gli intellettuali nostrani vi si sono trovati coinvolti come ad una fiera e non sempre con dignità. In senso lato fanno parte anche loro di quella «intellighenzia artificiale» di cui Weizenbaum parla nel libro, un gruppo di studiosi e ricercatori talmente coinvolti nel destino dell'informatica da essere pronti a giurare sulle sue proprietà nascoste e a venire.

Il libro di Weizenbaum si propone di ricostruire e di spiegare cosa i computer veramente sono, cosa possono fare, cosa non possono (in base alle conoscenze presenti) promettere o cosa, per la loro natura infine, non «debbono proprio fare». Qui non «debbono» ha il carattere di quella decisione «democratica» sulla scienza di cui parla P.K. Feyerabend. La discrezionalità rispetto ai mezzi tecnologici significa che alcuni mezzi non sono «appropriati» a certi fini ed una loro applicazione in questo caso porta a distorsioni, riduzioni, mistificazioni dei problemi e delle loro soluzioni. Weizenbaum dirà che ci sono problemi umani, sociali, politici che non possono essere affrontati con un GPS (*General Problem Solving*) proprio dell'Intelligenza Artificiale, perché una loro simulazione ne falsifica l'essenza e l'irriducibile complessità. Tutto il resto del campo è praticabile a patto che la Computer Science e l'Intelligenza Artificiale escano dal Sogno, dalla proiezione al futuro, ad un futuro di inimmaginabile potere, di un presente stato dell'arte che invece dovrebbe indurre ad essere realisti. Insomma una *Artificial Intelligence* non come vorrebbe essere, ma com'è.

Di questo *intelligence dream* Weizenbaum ci racconta la storia, che per molti versi corrisponde alla biografia scientifica.

Emigrato insieme ai genitori ebrei dalla Germania delle leggi antirazziali, Weizenbaum partecipa all'entusiasmo di coloro che negli Stati Uniti cercano di mettere a punto negli anni '50 le possibilità delle macchine di Turing.

La ricerca sui «cervelli elettronici» è affascinante: i computer non sono soltanto delle macchine calcolatrici. Sono delle macchine simulatrici. Esse rappresentano ed articolano modelli per simulare fatti e problemi del mondo «esterno». La loro grande novità, quella che il signore di fronte a me in treno sottovaluta, sta appunto nell'essere molto di più di una elaborata macchina da scrivere o di una calcolatrice tascabile. I computer, come dirà spesso Weizenbaum, sono «metafore». Con essi fin dall'inizio si è cercato di rappresentare un procedimento di risoluzione dei problemi. A questa impresa Weizenbaum ha partecipato creando uno dei primi programmi in cui un computer non solo sembrava «capire», ma poneva delle domande all'operatore. Il programma Eliza «ascoltava» come se fosse uno psicoterapista e conosceva il modo di arricchire il proprio linguaggio, rubando all'operatore spezzoni o frammenti di frasi. In più «sapeva» che se l'operatore usava frasi del tipo «Tutti ce l'hanno con me», doveva porgli una domanda che cominciava per «Chi in particolare ce l'ha con te?».

Questa simulazione inventata da Weizenbaum rivela ben presto tutto il suo ambiguo fascino. L'operatore «prende» la simulazione come realtà e «proietta» sul computer più di quanto la simulazione richieda. Il mondo scientifico ed ufficiale intorno alla rivista «Psichology» saluta Eliza come la possibilità di «democratizzare» e rendere accessibile a tutti la psicoterapia. Weizenbaum si ribella a questa interpretazione e fa notare che è semplicemente assurdo compiere una simile proiezione. Si accorge di aver aperto non una nuova fase tecnica, ma di avere aperto la porta del simbolismo al computer. «Il programma» della macchina diventa affidabile proprio perché garantito dalla co-

stanza e logicità non affetta da emozioni, neutra, e perché allo stesso tempo capace di simulare una situazione di relazione dialogica.

«Eliza creò la straordinaria illusione di aver capito la mente di molte persone con cui aveva conversato. Dimenticavano ciò che sapevano perfettamente (di conversare con una macchina) come gli spettatori a teatro dimenticano che l'azione a cui si assiste non è «reale».

Eliza come l'Eliza Doolittle di Pigmalione suggestionava l'interlocutore, «anche se non era chiaro, come per Miss Doolittle, se diventasse più o meno intelligente».

Il computer non valeva per quello che offriva in quel momento, ma per la potenza di metafora che rappresentava. Come era già successo con l'orologio, la nuova invenzione non era una macchina per segnare il ricorrere ciclico di certi avvenimenti naturali, ma il modello, in quanto macchina autoregolantesi, per descrivere gli organismi naturali e l'uomo stesso. Così la filosofia contemporanea se ne servì per costruire l'immagine dell'uomo e del suo funzionamento e altrettanto fecero la fisiologia e la biologia.

Allo stesso modo nel 1958 due colleghi di Weizenbaum dichiarano il loro credo: «Si tratta di creare macchine che pensano, imparano, creano e la cui abilità nel fare queste cose aumenti fino a che il campo dei problemi che sono in grado di trattare abbia estensione paragonabile a quella del campo a cui è stata applicata la mente umana». È il credo di quelli che Lous Fein definisce, prima di Weizenbaum, l'«intellighenzia artificiale».

L'ambiguità del campo sta nel sostituire continuamente ciò che non si conosce ancora con una metafora meccanica. Il gioco allora consiste nel definire termini come «pensare», «imparare», «capire», «creare», «mente» in base alla logica del computer e non in base a quello che il computer intende simulare. Ogni dieci anni l'«intellighenzia artificiale» promette, ad esempio, di essere prossima alla decodificazione del «linguaggio naturale» ed ogni dieci anni viene smentita: alla prova dei fatti il linguaggio naturale sfugge ad una codificazione, a meno che non lo si riduca nella sostanza. Come sarti che non sono a conoscenza dei corpi che devono vestire, gli informatici sostituiscono i propri modelli ai corpi e poi sono costretti a ridimensionare idealmente o con le forbici i corpi che non sono appropriati. Così l'intelligenza artificiale si pone come aiuto alla comprensione dei meccanismi dell'intelligenza umana, ma nella maggioranza dei casi compie solo un'opera di trivializzazione di un contesto e di un oggetto che pretende di assimilare alla propria invenzione.

C'è da chiedersi se Eliza fosse poi tanto casuale o se Weizenbaum non avesse voluto anzitutto provocare l'informatica nel suo punto d'onore.

Il testo che finalmente il pubblico italiano può apprezzare nella sua interezza delinea la storia, il funzionamento, le motivazioni ed il dibattito di quasi trenta anni di società del computer. Se sono accadute dall'84 ad oggi delle cose nuove (ma poi non tante e non così rilevanti) la sostanza del discorso rimane la stessa: come ogni scienza la «Computer Science» ha un'origine, e delle pretese implicite o esplicite. Bisogna affrontarle e parlarne perché il lavoro di ricerca e di invenzione rimanga onesto, cioè accetti la verifica di confronti esterni e non faccia prevalere le sue pretese di immagine sulle sue ipotesi scientifiche.

Di fronte all'affermazione di Simon: «Tutti gli uomini sono sistemi di elaborazione dell'informazione; si possono applicare ai loro affari i metodi principali della General Problem Solving (l'analisi euristica mezzi-fini)», Weizenbaum può anche consentire, ma poi domanda che la *Artificial Intelligence* definisca cosa significa per la mente umana «informazione» e cosa significhi «risoluzione di problemi».

Con una puntualità che gli viene rimproverata Weizenbaum tira fuori nel libro tutti i casi della vita politica americana in cui il *problem solving* è servito a ridurre e a mistificare questioni e tensioni sociali e politiche. Egli ricorda che i computer non sono strumenti neutri: essi riducono la realtà, la storia, il linguaggio, i conflitti al proprio funzionamento. L'oscenità non consiste in questo, se se ne ha la coscienza, ma nella pretesa secondo cui: «I modelli per il computer differiscono da quelli mentali perché i primi sono specificativi in modo esplicito. Le notazioni «matematiche», usate per descrivere il modello sono prive di ambiguità. È un linguaggio più chiaro e preciso dei linguaggi parlati, come l'inglese o il francese. Il linguaggio dei modelli per computer è più semplice. Il suo vantaggio sta nella chiarezza di significato e nella semplicità della sintassi. Il linguaggio di un modello per computer può essere compreso quasi da chiunque, indipendentemente dal grado di istruzione. Per di più qualsiasi concetto e connessione che possano essere tradotti in modo chiaro nel linguaggio ordinario possono venir tradotti nel linguaggio della modellistica dei computer».

Se i computer servono a «semplificare la vita», tra essi ed un sistema di pensiero totalitario c'è ben poca differenza. Weizenbaum ce l'ha contro i computer innalzati a visione del mondo e della conoscenza e ha ragione a metterci in guardia contro il totalitarismo già diffuso (dagli anni '60 almeno) in buona parte delle applicazioni a vasta scala di sistemi computerizzati, dal sistema di Difesa degli Stati Uniti, al sistema di controllo delle informazioni a livello mondiale (basti ricordare le denunce dell'Unesco ed il rapporto Mc Bride a questo proposito). Sarebbe stupido però dire che questo è un discorso contro i computer e contro l'Intelligenza Artificiale. Anzitutto perché Weizenbaum rimane un informatico e poi perché qui si tratta di una metafisica della macchina e delle sue applicazioni e si cerca di spiegare cosa invece è la fisica della stessa macchina.

Come dice Weizenbaum: «Newell, Simon, Shank, Winograd, per citare i maggiori rappresentanti della "intellighenzia artificiale" confondono la natura dei problemi che credono di "stare risolvendo". Costruiscono macchine per compiti specifici, come gli artigiani del XVII secolo costruivano macchine per tenere conto del tempo, per lanciare palle di ferro a distanze notevoli. Questi artigiani si sarebbero dolorosamente ingannati se si fossero lasciati portare dai loro successi alla conclusione di aver cominciato ad avvicinarsi ad una comprensione generale dell'universo, o anche alla conclusione che, siccome le loro macchine funzionavano, avevano dimostrato la validità dell'idea che le leggi dell'universo siano esprimibili in termini matematici. La *hubris* della intelligenzia artificiale si mostra proprio nella sua ricorrente adesione alle idee sbagliate sulle macchine che è riuscita a costruire».

Insomma le pretese generalizzatrici della Intelligenza Artificiale sembrano ignorare o voler rimuovere il processo di relativizzazione compiuto dalle scienze fisiche e biologiche negli ultimi anni. Pompate da un orgoglio tecnologico sembra che queste pretese possano anche fare a meno delle riflessioni su un secolo di studi sulla psiche, sulla parola, sui meccanismi della mente. Testi come *L'Io della mente* di Douglas Hofstaeder, o il suo stesso *Gödel, Escher, Bach*, suggeriscono una visione unificante di tutti gli organismi ed i processi nella ridefinizione di una «mente» o di un «io» a partire dagli stimoli della logica dei computer. La proposta può apparire affascinante, si può pensare che anche Gregory Bateson abbia contribuito ad un «allargamento» del concetto di «mente», invitando a trovare il filo che unisce la logica della natura alla logica della psiche. Eppure, fatta salva la discrezione di Bateson ed il suo rifuggire da generalizzazioni, l'impressione è quella di una facile riduzione dei fenomeni ai loro nomi. Si può sostituire all'onestà di una ricerca aperta la comodità di una definizione: così «intelligenza» e «mente» possono nascondere l'imbarazzo di un oggetto non facilmente afferrabile o peggio, come suggerisce Weizenbaum, il prendere come punto di riferimento della definizione il modello convenzionale, la macchina. L'operazione non è dissimile da quella di colui che definisce la circolazione del sangue nel corpo umano in base alle leggi di circolazione del traffico. (In effetti Harvey, lo scopritore della circolazione del sangue, paragonò il cuore ad una pompa idraulica).

Ci sembra che la posta in gioco non sia irrilevante. Certo, il dibattito ora si sta svegliando e si potrebbe ringraziare proprio l'informatica di avere rimesso al centro dell'attenzione i processi cognitivi e la natura stessa del soggetto della conoscenza. Ma, come mette bene in guardia Searle, ogni riduzione nel campo di definizione della conoscenza ha conseguenze non solo operazionali, ma propriamente morali. Se l'Intelligenza Artificiale vuol essere presa sul serio, e come dimostra Weizenbaum, lo vuole, perché pretende di essere lo specchio (lo specchio di questo secolo) della conoscenza, non può sottrarsi alla sua moralità, per una disciplina scientifica significa non barare sulle proprie ipotesi e sui propri strumenti.

PREFAZIONE
alla prima edizione

Questo libro è solo nominalmente un libro sui computer. Da un punto di vista significativo, il computer qui è usato come veicolo per far muovere certe idee che sono molto più importanti dei computer. Il lettore che dà un'occhiata a qualche pagina di questo libro e scappa terrorizzato per aver visto qui e là un'equazione o un poco di gergo informatico, dovrebbe ripensarci. Forse egli pensa di non sapere niente dei computer e, anzi, crede che i computer siano troppo complicati da capire per le persone normali. Ma una delle tesi principali di questo libro è proprio che noi, tutti noi, abbiamo trasformato il mondo più del lecito in un computer, e che questo rifacimento del mondo a immagine del computer è incominciato molto prima che nascessero i primi calcolatori elettronici. Ora che abbiamo i computer, diventa un po' più facile vedere questa trasformazione immaginaria che abbiamo attuato sul mondo. Ora possiamo usare il computer stesso — vale a dire, l'idea del computer — come una metafora che ci aiuti a capire quello che abbiamo fatto e quello che stiamo facendo.

Siamo tutti abituati a sentir dire che il computer è un nuovo, potente strumento. Ma poche persone hanno una qualche idea di quale sia l'origine della potenza di un computer. I capitoli da 1 a 3 sono dedicati proprio a spiegare questo. Con un modesto investimento di tempo ed energia intellettuale, chiunque sia in grado di leggere questa Prefazione dovrebbe potersela cavare anche con quei capitoli. I capitoli 2 e 3 saranno i più difficili, ma il lettore che non riesce a padroneggiarne il contenuto non dovrebbe per questo abbandonare il resto del libro. In realtà, l'unica tesi dei capitoli 2 e 3 è che i computer sono in un certo senso macchine «universali», e che possono (come verrà lì spiegato) fare «qualsiasi cosa». Il lettore che accetta questa affermazione sulla fiducia può benissimo decidere di saltare dal capitolo 1 (che dovrebbe leggere) al capitolo 4. Magari, dopo aver finito il libro sarà tentato di provare di nuovo a cimentarsi con i capitoli 2 e 3.

Il resto del libro contiene due tesi principali, che sono sostanzialmente: primo, che c'è differenza tra uomo e macchina, e secondo, che ci sono dei compiti che non *devono* essere affidati ai computer, indipendentemente dal fatto che sia *possibile* o meno affidarglieli.

La stesura di questo libro è stata per me un'avventura. In primo luogo, e cosa più importante, sono stato aiutato più di quanto io sappia dire dalla generosità e dal sostegno intellettuale ed emotivo datomi da persone che non mi devono assolutamente nulla. Ora sono io a dover loro moltissimo. Penso principalmente a Lewis Mumford, quel grande vecchio, a Noam Chomsky e a Steven Marcus, il critico letterario. Ciascuno di loro ha letto gran parte del manoscritto in preparazione (Lewis Mumford l'ha letto tutto), contribuendo con critiche molto sagge e utili. Ma in più, ciascuno di loro mi ha incoraggiato ad andare avanti, quando disperavo di farcela. Perché ho avuto spesso ragione di disperare. Sono profondamente consapevole, per esempio, del fatto che non c'è nulla di quanto dico in questo libro che non sia stato detto meglio, e certamente con migliore eloquenza, da altri. Ma, come i miei amici non hanno smesso di farmi notare, è sembrato importante ripetere più e più volte queste cose. E, come ha osservato spesso Lewis Mumford, talvolta è importante che sia un membro dell'*establishment* scientifico a dire alcune cose che gli umanisti hanno gridato migliaia di volte.

Più formalmente, sono debitore alla mia Università, il Massachusetts Institute of Technology, per avermi concesso un congedo di due anni che mi ha permesso di dedicarmi prima alla riflessione che ha preceduto la scrittura, poi alla scrittura stessa. Ho passato il primo di quei due anni felici al Centro di studi avanzati nelle scienze comportamentali di Stanford, in California [1]. È lì che ho incontrato Steven Marcus, come anche altri membri del Centro, che hanno lottato strenuamente per istruire un primitivo ingegnere come me. I nomi di John Platt, Paul Armer, Herbert Weiner, Fredrick Redlich, Alexander Mitscherlich e Israel Scheffer mi vengono immediatamente in mente. Li ringrazio per i loro sforzi. Nessuna mia mancanza, però, va imputata ad alcuno di loro. Ho passato il secondo anno come *Vinton Hayes Research Fellow* all'Università di Harvard. Lì ho avuto la fortuna di poter riallacciare una collaborazione nata ai tempi dell'università, quella con il professor Thomas Cheatham, un bravissimo scienziato dei computer. Egli si è preso la briga di leggere quasi l'intero manoscritto così com'era uscito dalla mia penna. Il professor Hilary Putnam del Dipartimento di filosofia di Harvard mi ha concesso molte ore del suo prezioso tempo. Senza il suo aiuto, il suo incoraggiamento e la sua guida sarei caduto in molte trappole in più. È stato anche un colpo di fortuna che Daniel C. Dennett, uno straordinario giovane filosofo della Tufts University, si sia trovato a passare un anno a Harvard proprio mentre c'ero anch'io. La sua pazienza con la mia ingenuità filosofica è stata illimitata. Non potrò mai ripagare adeguatamente il mio debito verso tutte queste persone.

Queste poche parole di ringraziamento servono a riconoscere il fatto che questo libro — come, sospetto, la maggior parte degli altri libri — ha molti coautori il cui nome non comparirà in copertina. Ma in questo caso tale am-

[1] La mia borsa di studio è stata finanziata in parte dalla Concessione n. SSH71-01834 A01 della National Science Foundation, dal Programma di Ricerca Applicata ai Bisogni Nazionali stipulato dalla NSF con il Centro. Naturalmente le opinioni, le tesi, le conclusioni e le raccomandazioni qui contenute sono interamente mie e non riflettono necessariamente le opinioni dei finanziatori.

missione sarebbe assai incompleta se non comprendesse un riconoscimento del contributo critico dato al manoscritto dal suo redattore, Aidan Kelly. Non posso riassumere in poche parole quello che egli ha fatto. Forse il lettore capirà se dico semplicemente che Aidan Kelly è un poeta.

Infine, chiunque abbia scritto un libro sa quale enorme fardello l'impresa imponga sulla famiglia dell'autore. Mia moglie Ruth ha sopportato i ritiri nel mio studio con la massima buona volontà e pazienza. Mi ha aiutato a superare le inevitabili battaglie contro il senso di colpa che schiaccia un autore quando scrive — perché in quei momenti egli non sta con la sua famiglia, pur essendo in casa — e quando non scrive — perché allora non sta facendo quello che si è proposto di fare. I miei bambini hanno contato le pagine che si ammucchiavano sul mio tavolo. E hanno sofferto quando, come è successo spesso, il mucchio di pagine nel cestino aumentava più rapidamente di quello sul tavolo. Soprattutto, essi hanno sopportato con serenità le relazioni senza fine sullo stato del lavoro che occupavano le conversazioni dell'ora di cena. Questo libro è di Ruth e dei nostri figli tanto quanto esso è mio.

Joseph Weizenbaum

Autunno 1975
Cambridge, Massachusetts

PREFAZIONE
all'edizione inglese del 1984

La pubblicazione di *Il potere del computer e la ragione umana* da parte dell'editore britannico Penguin mi offre l'opportunità e l'occasione di pensare al tempo passato da quando l'edizione originale americana venne pubblicata da Freeman and Company nel 1976.

Una domanda da farsi è se il libro abbia in qualche modo reso diverso il mondo. Ho pensato spesso che ora si sentono molte meno predizioni e vanterie euforiche da parte della comunità che lavora sull'«intelligenza artificiale» (in seguito, useremo spesso per comodità l'abbreviazione *IA*, NdT) di quante se ne sentissero prima della pubblicazione del libro. Forse questo è stato un effetto delle mie critiche alla comunità dell'*IA*. D'altra parte, forse l'*IA* è soltanto maturata un poco. Con più sicurezza direi che il libro ha incoraggiato molte persone a credere nelle proprie idee, e che questo rimane il suo merito principale. È davvero straordinariamente difficile, per la maggior parte delle persone, tener duro su certe idee nell'assenza di prove credibili che esse siano condivise da altri. Ma il sapere che anche una singola altra persona ha un pensiero che ho io conferma e convalida quel pensiero. Da questo punto di vista c'è una differenza molto più importante tra una persona e due che tra due e, mettiamo, dieci. Questo lo so proprio grazie alle molte lettere di lettori che ho ricevuto in questi anni, lettere che esprimono la gratitudine per aver visto, in un certo senso, confermate le loro ragioni. Sembra che siamo arrivati a un momento in cui l'idea che vi siano differenze tra gli esseri umani e le macchine, che vi siano esperienze che gli esseri umani possono fare ma le macchine no, e che perciò (almeno!) ci sono pensieri che gli uomini ma non le macchine possono avere, ebbene siamo arrivati a un momento in cui la concezione di queste idee è una preoccupazione solitaria, una faccenda che fa dubitare chi le concepisce della propria sanità mentale. Nel libro ho scritto che il fatto stesso che sia necessario discutere di queste idee è un indice della follia del nostro tempo. Il pericolo ora è che il dibattito si spenga; non perché sia stata infine universalmente riconosciuta l'assurdità di considerare gli esseri umani equivalenti a dei computer, ma perché le voci a difesa delle posizioni antropocentriche sono diventate ancora più flebili.

Molte cose sono successe nel mondo dalla prima apparizione di *Il potere del computer e la ragione umana* nelle librerie americane nel 1976. Penso che ci sia accordo universale sul fatto che il mondo è diventato un posto molto più pericoloso in questo frattempo. Nel complesso, le soluzioni tecnologiche che venivano pubblicizzate come strumenti che ci avrebbero dato una maggiore sicurezza e stabilità internazionale, hanno in realtà contribuito principalmente alla destabilizzazione e a una maggiore insicurezza. Lo stimolo delle esigenze militari, tra gli altri fattori, ha prodotto componenti elettronici sempre più piccoli e tecniche di produzione in serie quasi miracolose, come per esempio la *VLSI*, integrazione su scala molto grande. Computer più piccoli dal punto di vista fisico, ma più potenti da quello funzionale, hanno permesso al «carico utile» dei missili — cioè alla loro porzione che distribuisce morte — di aumentare, «migliorando» allo stesso tempo la loro precisione nel colpire il bersaglio. (Il progettista dei missili V-2 tedeschi, Werner von Braun, ha intitolato la sua autobiografia *Io miro alle stelle*. I londinesi, tuttavia, ricorderanno che di solito egli colpiva Londra. Le vittime della guerra moderna possono almeno avere la soddisfazione di sapere che le bombe che distruggono le loro città erano proprio destinate a distruggerle, e non ad altro). Una ricaduta accidentale di questo «progresso» è che i computer in generale sono diventati più piccoli, più veloci, più potenti e più economici. È questo effetto collaterale che viene strombazzato dai mezzi di comunicazione, e domina l'attenzione dell'opinione pubblica.

Sicuramente, ho mancato di prevedere la quasi unanimità e le enormi dimensioni dell'entusiasmo con cui il pubblico — almeno negli Stati Uniti — ha accolto il computer nelle faccende quotidiane. I portafogli dei cittadini sono rigonfi di un numero sempre maggiore di tessere di plastica con striscia magnetizzata, ciascuna delle quali, venendo letta da una macchina, dà al portatore l'accesso a qualche attività di routine — magari il transito su un mezzo di trasporto pubblico, o il ritiro di denaro da uno sportello bancario automatico — che fino a poco tempo fa era mediata da altre persone. La maggior parte dei cittadini degli Stati Uniti oggi possiede computer di qualche tipo. Microcomputer abitano lavatrici e macchine fotografiche, motori di automobile e orologi, macchine per scrivere e telefoni. I più tra noi sono tanto inconsapevoli della presenza di questi «micro» quanto di quella dei molti motorini elettrici che ci circondano. Abbondano anche i calcolatori elettronici tascabili, per lo più molto economici. Tuttavia, il computer da casa completamente sviluppato, vale a dire una macchina indipendente, completa, programmabile, dotata dei suoi apparecchi di input/output, benché sia disponibile e venga usato diffusamente nelle attività commerciali, non è ancora diventato un elettrodomestico onnipresente come, mettiamo, il telefono.

I tipi di computer che si sono fatti strada nella coscienza dell'opinione pubblica sono quelli programmati per giocare nelle sale-giochi e quelli da collegare agli apparecchi televisivi. I giochi computerizzati esistevano certamente già all'epoca in cui scrissi il libro. *A posteriori*, risulta ovvio che proprio quelli che erano più popolari al MIT e nelle altre università, per esempio il gioco «Guerra spaziale», erano i precursori dei *videogame* che dominano oggi nelle

sale-giochi e nelle case. Non mi è affatto chiaro perché io non ne abbia afferrato il significato otto anni fa.

Secondo me è fuori discussione, oggi, che il fenomeno dei *videogame*, almeno nella veste che ha negli Stati Uniti, conferma prepotentemente il pessimismo culturale generale che *Il potere del computer e la ragione umana* esprime. Per quanto disperante sia vedere l'uso che la nostra società fa della televisione, tale disperazione è ingigantita dalla visione di innumerevoli ragazzi in piedi, ipnotizzati, davanti allo schermo del computer, con le mani che si muovono come quelle di un soldato rintronato dall'esplosione di una bomba. Io penso che quello che sta succedendo ai giovani nelle sale-giochi e in alcune aule scolastiche sia una parabola del nostro tempo, una storia triste e che turba.

Per capire il contenuto della maggior parte dei videogiochi, è sufficiente esaminare un campione dei più seguiti programmi televisivi (almeno negli Stati Uniti) per rendersi conto che esso consiste per lo più in quello che in tedesco si chiamerebbe *Unsinn*, *Blödsinn*, *Wahnsinn*, vale a dire nonsenso, stupidità e follia. Sciocche cosiddette *situation comedies* sono l'unica alternativa a un flusso altrimenti costante di violenza sfrenata. Al sabato e alla domenica mattina viene ripresentato lo stesso messaggio, ora però sotto forma di cartoni animati, in modo che i bambini possano assorbirlo in modo ancora più diretto. La massa dei videogiochi presenta esattamente lo stesso messaggio, tradotto, naturalmente, in modo da avvantaggiarsi del nuovo mezzo. Tuttavia, laddove lo spettatore televisivo riceve passivamente, il giocatore di *videogame* partecipa attivamente. In termini pratici questo significa che, mentre lo spettatore televisivo guarda, diciamo, il comandante del sottomarino lanciare siluri contro le navi «nemiche» e gridare di gioia quando i suoi bersagli scompaiono tra le onde, il giocatore di *videogame* lancia i siluri e prova egli stesso l'ebbrezza della corsa del siluro. Stavo quasi per scrivere «l'ebbrezza di *uccidere*», e dietro a questo quasi-scivolone della penna si nasconde un punto importante: oso dire che pochissimi comandanti di sottomarini, quelli veri, fanno l'esperienza dell'*uccisione* quando schiacciano il bottone che dà inizio alla corsa del siluro verso il bersaglio, e lo stesso vale per i soldati sui bombardieri, quando lanciano le loro bombe. Per la maggior parte, gli esseri umani sarebbero incapaci di azioni come queste, se non avessero la possibilità di rimanere a quella che un medico chiamerebbe «distanza clinica» dalle conseguenze finali delle loro azioni. Un modo meno eufemistico per dire la stessa cosa è che ci vuole un addestramento molto intenso, a base di stordimento psichico, prima che una persona normale possa lanciare un siluro che affonderà una nave, o sganciare una bomba che dilanierà la gente parecchie miglia più in basso.

Le accademie militari e navali di tutto il mondo dedicano infatti molto tempo ed energia a periodi di addestramento in cui tutte le operazioni tecniche, tranne le più innocue e apparentemente innocenti, sono completamente slegate dall'atto dell'uccidere. Considerazioni simili valgono per coloro che producono componenti di sistemi d'arma, o sistemi sempre più intelligenti per guidare infallibilmente i missili sui bersagli cui i loro padroni li hanno destinati. In questo mondo c'è un mercato pronto ad accogliere a braccia aperte persone che siano già psichicamente intorpidite quando entrano a far parte della forza-

lavoro, vale a dire, che siano stupendamente allenate a non interessarsi della connessione tra ciò che fanno e l'effetto finale che ciò che fanno ha su quelli che potremmo chiamare «utenti ultimi» del prodotto del loro lavoro, in altre parole, le loro vittime finali.

La maggior parte dei videogiochi che io ho visto servono ad allenare proprio questa capacità. Astronavi e aeroplani vengono abbattuti in gran numero. Bombe da molti megaton cadono a pioggia su intere nazioni. Molti giochi sono organizzati in modo tale da non lasciare superstiti: «vincere» equivale allora a riuscire a far andare avanti il gioco più a lungo degli altri giocatori, il che in generale significa essere un *killer* più rapido ed efficace. Sul mercato americano c'è addirittura un videogioco chiamato «La rivincita di Custer», dove il vincitore è colui che è riuscito a *stuprare* il maggior numero di donne indiane! In quel gioco, l'offesa più crudele e spaventosa che un uomo può portare a una donna viene ripulita di tutto il tormento, l'orrore, l'angoscia. Restano soltanto operazioni astratte su pulsanti di plastica. Questo è ciò che intendo per stordimento psichico.

Nel momento in cui scrivo, le istituzioni scolastiche americane sono in uno stato di quasi-panico. I *mass-media* hanno stabilito che siamo la «Società dell'informatica». La pressione sulle scuole perché usino il computer nell'istruzione dei bambini di cinque o più anni è quasi irresistibile. Gli Stati Uniti sembrano essere in preda a una moda che ricorda le radio CB, il cubo di Rubik, l'*hula-hoop* e le altre follie del passato. Gli effetti della malattia di oggi, però, probabilmente non saranno così banali come quelli del passato. La mente dei nostri giovani verrà plasmata in modi nuovi. Questo non è semplice intrattenimento. C'è un legame tra questo e il fenomeno dei videogiochi che ho appena discusso? Sì, ed è almeno di due tipi.

In primo luogo, c'è il fatto che gran parte dei *personal computer* più piccoli, cioè di quelli destinati all'uso diretto da parte degli studenti, già installati nelle scuole, sono predisposti per giocare i più popolari *videogame*. Per molti giovani, questi giochi sono proprio ciò che rende il computer attraente. In altre parole, possiamo aspettarci che l'«educazione» alla violenza senza sensi di colpa e allo stordimento mentale, l'addestramento fornito dalle sale-giochi e spesso anche dai computer domestici, venga continuata e rafforzata dalla scuola, anche in virtù di quel po' di autorità morale che alla scuola è rimasta.

In secondo luogo, c'è il fatto che gran parte del normale lavoro di calcolo che verrà insegnato ai bambini al fine di prepararli per il loro ruolo nella Società dell'informatica, cose come la costruzione di modelli, per esempio, ha una capacità di separare quello che uno fa dagli effetti di quello che uno fa paragonabile a quella dei videogiochi che ho appena discusso; soltanto che il processo è più sottile. Certamente, è relativamente facile evitare di contaminare le menti dei giovani, pur insegnando loro a progettare e usare i modelli al computer. Basta avere insegnanti attenti, e che capiscano essi per primi che cosa i modelli sono e cosa si possa e non si possa imparare dai modelli.

Nel libro ho discusso modelli e teorie (capitolo 5). Ciò che è importante nel contesto attuale è il fatto che i modelli incarnano soltanto le caratteristiche *essenziali* di ciò che intendono rappresentare. Se un modello di un'automobile

è destinato alle prove nella galleria del vento, allora è importante la forma esterna del modello, ma in esso non c'è bisogno né di sedili né di alcun altro arredamento interno che sarà presente nella vera automobile. La scelta degli aspetti della realtà che il modello deve o non deve incarnare dipende interamente dagli scopi di chi costruisce il modello. Ma qualunque sia lo scopo, un modello — e qui penso in particolare ai modelli computerizzati di aspetti della realtà — deve necessariamente *lasciare da parte quasi tutto* ciò che è in realtà presente nell'oggetto vero e proprio. Quindi, ciò che si può imparare manipolando il modello è *in ogni caso* rigorosamente limitato. Chiunque sappia questo e ne riconosca l'importanza, e lo tenga presente nell'insegnare a usare il computer, ha una possibilità di immunizzare i suoi studenti dal rischio di credere troppo, di dare eccessiva importanza ai risultati del loro lavoro con il computer.

In realtà, tuttavia, pochi insegnanti (a qualsiasi livello!) conoscono e tengono in giusto conto limitazioni di questo tipo. E poca, o nessuna, enfasi viene data a questi argomenti nella maggior parte dei corsi di laurea in informatica. Discorsi come questi sono ritenuti *solamente* filosofici. Inoltre, la fretta di inondare le scuole con i computer impedisce un aggiornamento adeguato degli insegnanti. La risposta delle istituzioni scolastiche americane all'ovvia impossibilità di preparare insegnanti in numero appropriato per le dimensioni che l'alluvione di computer sulle scuole sta avendo consiste nel trasformare il disastro in un trionfo: di conseguenza, ci si *vanta* del fatto che in generale gli insegnanti americani di materie che hanno a che fare col computer sanno sulla loro materia meno di quanto sappiano i giovani che è loro compito istruire.

Un risultato inevitabile di questa situazione è, temo, che quasi tutti i giovani che ricevono la loro introduzione formale ai computer nelle scuole elementari e medie si ritroveranno dotati di una certa capacità di creare modelli al computer, ma privi di qualsiasi conoscenza sui loro limiti. Cresceremo così nuove generazioni di giovani del tutto preparati a eseguire sempre più rattoppi tecnologici su un mondo che ne è già strapieno, saturo, intossicato. Ma ciò che preoccupa anche di più è che il tipo di semplicismo ingenuo di cui sto parlando qui ha come risultato, anzi è, una abdicazione alle responsabilità, una chiusura della mente alla realtà non accompagnata da un senso di incompletezza; in altre parole, un intorpidimento psichico di fronte alle conseguenze di ciò che si fa, analogo a quello prodotto dalle sale-giochi.

Poco fa ho evocato la televisione come modello per il contenuto dei videogiochi. Il paragone può essere esteso. A me sembra che non si possa determinare con precisione, e neppure stimare, quale sia l'effetto della televisione sull'ascoltatore, se si limita l'analisi ad ascoltatori singoli davanti al loro apparecchio. L'effetto della televisione su una società è in gran parte determinato dal fenomeno di milioni di menti continuamente e simultaneamente innaffiate da ciò che la televisione trasmette. È lo stesso con i computer. L'effetto del giocare, mettiamo, il videogame «Defender», in cui alcune città sono oggetto di bombardamento nucleare, sarebbe diverso se a giocare fosse una sola persona, e isolata, rispetto a quello che è se si gioca in un contesto già permeato dallo spirito, dall'ideologia di quello e di molti altri giochi, che sono stati tutti profondamente interiorizzati dai coetanei di chi gioca. L'effetto sui bam-

bini dell'esposizione all'idolatria del tipo di razionalità richiesto (almeno per ora) per poter programmare i computer, dell'ulteriore elevazione del calcolo al di sopra del giudizio, va compreso nel contesto della cultura in cui ha luogo. La nostra cultura, come ho cercato di dire nel libro, ha già quasi stabilito che l'unica conoscenza legittima che si può avere del nostro mondo sia quella scientifica. Tutto il pensiero, tutti i sogni, tutte le sensazioni, tutte le altre fonti di comprensione sono già state delegittimate. L'indottrinamento della mente dei nostri bambini con un'idolatria dei computer semplicistica e disinformata — ed è quasi certamente questo che il computer nella scuola è e sarà — è un fenomeno pandemico. Basta questo, secondo me, per convalidare il pessimismo culturale espresso nell'edizione originale del libro.

Naturalmente, sarei stato felice di salutare il lettore inglese con il messaggio che gli avvenimenti successivi alla pubblicazione del libro nel 1976 hanno dimostrato false le ragioni di disperazione che caratterizzano alcune delle sue pagine. Forse la situazione in Gran Bretagna è di maggior civiltà rispetto agli Stati Uniti. Se è così, considerate questo libro un monito.

Joseph Weizenbaum

1983
Cambridge, Massachusetts

INTRODUZIONE

Nel 1935 Michael Polanyi, all'epoca titolare della cattedra di Chimica fisica alla Victoria University di Manchester, in Inghilterra, si ritrovò in modo improvviso e sconvolgente a fronteggiare certe questioni filosofiche che da allora in poi avrebbero dominato la sua vita. All'origine di questo evento traumatico c'era Nicolai Bukharin, uno dei più importanti teorici del Partito comunista sovietico, il quale disse a Polanyi che «nella società socialista il concetto di scienza come obiettivo fine a se stesso scompare, perché l'interesse degli scienziati tende spontaneamente a rivolgersi ai problemi dell'attuale Piano quinquennale». Polanyi si rese conto che la «mentalità scientifica aveva evidentemente prodotto una concezione meccanicistica dell'uomo e della storia in cui non restava spazio per la scienza stessa». E inoltre, che «questa concezione negava assolutamente che il pensiero potesse avere un qualche potere intrinseco, e negava così l'esistenza di buone ragioni per rivendicare la libertà di pensiero» [1].

Io non so quanto tempo Polanyi abbia pensato di dedicare all'elaborazione di una tesi favorevole a una diversa visione dell'uomo e della storia. Il fatto di essere rimasto sconvolto già di per sé mostra che egli era in profondo disaccordo con Bukharin, e che quindi egli aveva già pensato all'uomo in termini diversi, anche se sul momento non era in grado di dare una forma esplicita a tale convinzione. Può darsi che egli abbia deciso di scrivere una risposta alla posizione di Bukharin, basandosi soltanto sulla sua esperienza di scienziato, e che l'abbia fatto immediatamente. Come si è visto, però, il confronto con la filosofia innescato dalla rivelazione di Bukharin avrebbe richiesto il completo impegno di Polanyi da allora fino ai giorni nostri.

Ho ricordato questo pezzetto di storia per due motivi. Il primo è illustrare come certe idee che a prima vista sembrano essere semplici e ovvie, e che dovrebbero perciò risultare universalmente credibili una volta espresse, finiscono talvolta per essere delle boe di segnalazione per le zone tempestose dei profondi mari intellettuali. L'idea che la scienza è creativa, che l'atto creativo nella scienza è equivalente a quello nell'arte, che soltanto individui autonomi

[1] M. Polanyi, *La conoscenza inespressa*, Armando, Roma 1979, pp. 3-4.

possono creare, è un'idea molto semplice e, si potrebbe pensare, ovvia. Eppure Polanyi, come molti altri, ha passato quasi una vita a esplorare la base su cui quell'idea è ancorata, e il turbolento mare di implicazioni che la circonda.

La seconda ragione per cui ho citato questa storia è che ho l'impressione di riviverla, in parte, personalmente. A me il trauma fu causato non da qualche importante personaggio politico che esponesse la sua filosofia della scienza, bensì da certe persone che hanno insistito nell'interpretare in modo non corretto uno dei miei lavori. Dico questo senza rancore, e certamente non per difendermi. Anzi, le interpretazioni cui mi riferisco tendevano semmai a sopravvalutare l'entità dei risultati che avevo ottenuto, e certamente la loro importanza. No, ricordo qui quel lavoro soltanto perché mi sembra che indichi la strada più parsimoniosa per identificare il tema che voglio discutere.

Quel lavoro l'ho scritto nel periodo 1964-1966, e venne pubblicato sulle riviste di scienza dei computer nel gennaio 1966 e nell'agosto 1967. Per riassumerlo brevemente, si può dire che avevo elaborato un programma per il computer che permetteva all'operatore di «conversare» in inglese con la macchina. L'interlocutore umano doveva comporre la sua parte di conversazione su una macchina per scrivere collegata a un computer e questo, controllato dal mio programma, analizzava il messaggio ricevuto, componeva una risposta in inglese e la faceva stampare dalla scrivente.

Scelsi il nome *Eliza* per il programma di analisi del linguaggio perché, come l'Elisa della leggenda di Pigmalione, esso poteva imparare a «parlare» sempre meglio. Siccome ogni conversazione deve avere un argomento, vale a dire deve avvenire all'interno di un contesto, il programma era organizzato su due piani: il primo conteneva l'analizzatore del linguaggio, il secondo un copione. Il copione è un insieme di regole simili a quelle che potrebbe usare un attore che deve improvvisare su un certo tema. In questo modo, *Eliza* poteva essere messo in grado di sostenere una conversazione su come si cucinano le uova, o sulla gestione di un conto corrente bancario, e così via. Ogni particolare copione permetteva ad *Eliza* di assumere un ruolo diverso nella conversazione.

Per il mio primo esperimento diedi ad *Eliza* un copione progettato per permetterle di recitare (dovrei in realtà dire «parodiare») la parte di uno psicoterapeuta rogeriano impegnato in un primo colloquio con un paziente. Lo psicoterapeuta rogeriano è relativamente facile da imitare, perché gran parte della sua tecnica consiste nel far parlare il paziente, ripetendogli le sue stesse affermazioni. Il seguente colloquio tra una giovane donna ed *Eliza* che recita da terapeuta illustra tanto la tecnica rogeriana, che incoraggia il paziente a continuare a parlare, quanto il funzionamento del programma *Eliza*. Prima a parlare è la giovane donna. Le risposte del computer sono stampate tutte in maiuscolo [2].

[2] Questa «*Conversazione*» è tratta da J. Weizenbaum, *ELIZA - A Computer Program For The Study of Natural Language Communication Between Man and Machine* («ELIZA: un programma per il computer per studiare la comunicazione in linguaggio naturale tra uomo e macchina»), in *Communications of the Association for Computing Machinery*, vol. 9, nr. 1 (gennaio 1965), pp. 36-45.

Gli uomini sono tutti uguali

IN CHE SENSO

Trovano sempre il modo di tormentarci

PUO' FARMI UN ESEMPIO

Beh, è il mio ragazzo che mi ha fatto venire qui

E' IL SUO RAGAZZO CHE L'HA FATTA VENIRE

Dice che sono sempre depressa

MI DISPIACE SENTIRE CHE È DEPRESSA

È vero. Sono infelice

PENSA CHE IL FATTO DI VENIRE QUI L'AIUTERÀ A NON ESSERE INFELICE

Ho bisogno di aiuto, questo mi sembra chiaro

COSA SIGNIFICHEREBBE PER LEI POTER RICEVERE UN AIUTO

Forse potrei imparare a tollerare mia madre

MI DICA DI PIÙ SULLA SUA FAMIGLIA

Mia madre si prende cura di me

CHI ALTRO NELLA SUA FAMIGLIA SI PRENDE CURA DI LEI

Mio padre

SUO PADRE

Lei è come mio padre, in certe cose

CHE COSA TROVA DI SIMILE

Lei non è aggressivo, ma penso che non voglia che io me ne accorga

CHE COSA LE FA PENSARE CHE IO NON SIA AGGRESSIVO

Non discute quello che dico

PERCHÉ PENSA CHE IO NON DISCUTA QUELLO CHE DICE

Perché ha paura di me

LE FA PIACERE PENSARE CHE HO PAURA DI LEI

Mio padre ha paura di tutti

COS'ALTRO LE VIENE IN MENTE QUANDO PENSA A SUO PADRE

Prepotenza

QUESTO HA QUALCOSA A CHE VEDERE COL FATTO CHE IL SUO RAGAZZO L'HA FATTA VENIRE QUI

Questo programma *Doctor* (con tale nome veniva chiamato *Eliza* quando recitava da psichiatra) divenne presto famoso al Massachusetts Institute of Technology, dove vide la luce, specialmente perché era un programma di cui era facile dare una dimostrazione. In genere, gli altri programmi non riuscivano a esibire altrettanto chiaramente il potere di elaborazione dei dati di un computer, se non a visitatori che avessero già qualche conoscenza specializzata, per esempio in qualche ramo della matematica. *Doctor*, invece, poteva essere apprezzato in certa misura da chiunque. La sua potenza come veicolo dimostrativo era ulteriormente rafforzata dal fatto che il visitatore poteva partecipare personalmente alla dimostrazione. Presto altre copie di *Doctor*, scritte basandosi sulla descrizione che ne avevo pubblicato, incominciarono ad apparire in altri istituti degli Stati Uniti. Il programma divenne noto su scala nazionale; in certi ambienti, anzi, divenne un giocattolo nazionale.

Lo *shock* che ho ricevuto quando *Doctor* è diventato famoso e ha cominciato a essere «giocato» da molta gente è da addebitare principalmente a tre avvenimenti.

1. Un buon numero di psichiatri credette seriamente che il programma per computer *Doctor* avrebbe potuto sfociare in una nuova forma, quasi completamente automatizzata, di terapia. Colby et. al., per esempio scrissero:

> Ci vorrà ancora del lavoro prima che il programma sia pronto per l'uso clinico. Se il metodo si dimostrasse benefico, rappresenterebbe uno strumento terapeutico facilmente accessibile per quegli ospedali e centri psichiatrici che soffrono di penuria di terapeuti. In virtù della capacità, propria dei computer moderni e di quelli futuri, di operare simultaneamente su più problemi, un sistema di computer opportunamente programmati potrebbe trattare centinaia di pazienti ogni ora. Il terapeuta umano, coinvolto nel progetto e nel controllo di questo sistema, non verrebbe sostituito, ma diventerebbe molto più efficace, perché i suoi sforzi non sarebbero più limitati al rapporto uno-a-uno tra paziente e terapeuta che esiste oggi [3].

[3] K.M. Colby, J.B. Watt, and J.P. Gilbert, *A Computer Method of Psychotherapy: Preliminary Communication* («Un metodo computerizzato per la psicoterapia: comunicazione preliminare»), *The Journal of Nervous and Mental Disease*, vol. 142, 1966, nr. 2, pp. 148-152. (E il dottor Colby non era solo in questo suo entusiasmo per la psicoterapia svolta dal computer. Carl Sagan, il noto astrofisico, commentò a proposito di *Eliza* (*Natural History*, vol. LXXXIV, nr. 1 gennaio 1975, p. 10): «Oggi non esiste un programma per computer che sia adeguato per l'uso psichiatrico, ma lo stesso si potrebbe dire di certi psicoterapeuti umani. In un'epoca in cui sembra

Avevo sempre creduto che, per poter aiutare qualcuno ad affrontare i suoi problemi emotivi, fosse indispensabile partecipare all'esperienza di quei problemi, e arrivare a capirli proprio grazie a questa loro individuazione empatica. Indubbiamente, ci sono molte tecniche per facilitare la proiezione immaginaria del terapeuta nella vita interiore del paziente. Ma che fosse possibile che anche un solo psichiatra praticante proponesse di soppiantare questa componente fondamentale del processo terapeutico con mezzi puramente tecnici, questo non l'avrei mai immaginato! Chissà cosa deve pensare del suo lavoro uno psichiatra capace di fare una proposta simile, e di credere che la più semplice parodia meccanica di una tecnica di colloquio possa afferrare l'essenza di un incontro tra esseri umani! Forse sono proprio Colby *et al.* a dare una spiegazione, quando scrivono:

> Un terapeuta umano può essere visto come un elaboratore di informazioni e un produttore di decisioni, dotato di un insieme di regole decisionali strettamente legate a obiettivi di breve e di ampio respiro |...| In queste decisioni, egli è guidato da rozze regole empiriche che gli suggeriscono cosa è opportuno dire e non dire in determinate situazioni. Incorporare queste procedure — al livello posseduto dal terapeuta umano — nel programma è un'impresa difficile, ma stiamo cercando di muoverci in questa direzione [4].

Che immagine può avere uno psichiatra del suo paziente, se considera se stesso non già un essere umano impegnato ad agire come guaritore, bensì un elaboratore di informazioni che segue certe regole ecc.?

Domande come queste mi fecero capire cosa intendesse Polanyi quando parlava di una «mentalità scientifica che aveva evidentemente prodotto una concezione meccanicistica dell'uomo».

2. Rimasi allibito nel vedere quanto rapidamente e profondamente le persone che conversavano con *Doctor* si lasciassero coinvolgere emotivamente dal computer, e come questo assumesse evidenti caratteri antropomorfici. Una volta la mia segretaria, che mi aveva visto lavorare al programma per molti mesi, e quindi certamente sapeva trattarsi soltanto di un programma per computer, incominciò a conversare con esso. Dopo pochi scambi di battute, mi chiese di uscire dalla stanza. Un'altra volta, dissi che avrei potuto modificare il sistema in modo da permettermi di registrare, e poi analizzare, tutte le conversazioni che avevano avuto luogo, mettiamo, durante la notte. Fui immediatamente bombardato di accuse di voler spiare i pensieri più intimi della gente; il che mostra

che siano sempre più le persone che, nella nostra società, hanno bisogno di consulenze psichiatriche, e in cui è sempre più diffuso l'uso di un computer da parte di più utenti, io posso immaginare lo sviluppo di una rete di terminali di computer psicoterapeutici, qualcosa di simile ad una serie di grandi cabine telefoniche in cui, per pochi dollari a sessione, sarebbe possibile parlare con uno psicoterapeuta attento, collaudato e in larga misura non autoritario»).

[4] *Ivi.*

in modo evidente che certe persone conversavano con il computer come se fosse una persona cui ci si poteva rivolgere per confidare i propri pensieri più intimi. Sapevo, naturalmente, che spesso tra uomini e macchine si creano legami emotivi di ogni tipo; è quello che succede, per esempio, con gli strumenti musicali, le motociclette, le automobili. E sapevo per lunga esperienza personale che i forti legami emotivi tra programmatori e il loro computer si formano spesso dopo un contatto anche breve con la macchina. Ciò di cui non mi ero reso conto era che un contatto estremamente breve con un programma per computer relativamente semplice potesse generare nelle persone normali delle enormi illusioni. Questa scoperta mi spinse ad attribuire una nuova importanza al problema del rapporto tra esseri umani e computer, e a decidere di pensarci su.

3. Un'altra reazione generale, e per me sorprendente, al programma *Eliza* fu la diffusione della credenza che esso costituisse la soluzione generale al problema della comprensione del linguaggio naturale da parte del computer. Nel mio lavoro, avevo cercato di dire che una soluzione generale a quel problema non esisteva, cioè che il linguaggio naturale può essere compreso soltanto nell'ambito di certi contesti; che anche gli esseri umani possono condividere questa visione contestuale solo fino a un certo punto, e quindi neppure loro sono incarnazioni di una soluzione generale di quel tipo. Ma queste conclusioni vennero spesso ignorate. In ogni caso, *Eliza* era un passo avanti così semplice e modesto! Il suo contributo, se ce n'era uno, consisteva nel sottolineare vivacemente ciò che molti altri avevano scoperto molto tempo prima: vale a dire, l'importanza del contesto per la comprensione del linguaggio. Il lavoro successivo di Winograd [5], molto più elegante e certamente molto più importante, è stato frainteso tanto quanto lo era stato *Eliza*. Questa reazione alla pubblicazione di *Eliza* mi mostrò, più chiaramente di quanto avessi visto fino allora, come un pubblico anche colto sia capace, e anzi desideroso, di attribuire proprietà enormemente esagerate a una nuova tecnologia che non capisce. Di sicuro, pensai, l'atteggiamento dell'opinione pubblica di fronte alle tecnologie emergenti dipende molto più dalle proprietà che vengono loro attribuite che da ciò che esse possono o non possono fare. Se, come sembra sia successo nei casi citati, le idee dell'opinione pubblica sono completamente infondate, ne consegue che le decisioni pubbliche saranno male indirizzate, e spesso sbagliate. Da queste considerazioni emergono delle domande difficili; quali precauzioni, per esempio, deve prendere uno scienziato prima di divulgare i suoi risultati? E di fronte a chi (o a che cosa) è responsabile?

Quando domande di questo tipo incominciarono a echeggiare in me, pensai — come forse aveva fatto Polanyi dopo il suo incontro con Bukharin — che era possibile liberarsi facilmente dei problemi e dei sospetti che mi si erano così prepotentemente presentati; magari con un breve e serio articolo. E in ef-

[5] T. Winograd, *Procedures as a representation for Data in a Computer Program for Understanding Natural Language* («Le procedure come rappresentazioni di dati in un programma per il computer per la comprensione del linguaggio naturale»), tesi di Ph. D. sottoposta al Dipartimento di matematica del MIT, 24 agosto 1970.

fetti scrissi un lavoro che toccava molti dei punti affrontati qui[6]. Ma, gradualmente, mi resi conto che certe questioni piuttosto fondamentali mi avevano contagiato in modo più cronico di quanto avessi pensato all'inizio. Probabilmente non me ne libererò mai.

Ci sono tanti modi diversi per affermare questi problemi fondamentali quanti sono i punti di partenza per affrontarli. In ultima analisi, l'oggetto della discussione è il ruolo dell'uomo nell'universo, niente di meno! Ma la mia preparazione professionale riguarda soltanto la scienza dei computer, il che equivale a dire (senza ironia) che sono assai poco colto; non ho né la competenza, né il coraggio e neppure la *chutzpah* (vocabolo yiddish che equivale più o meno a «faccia tosta», NdT) per scrivere di argomenti così grandiosi. Affronterò quindi problemi più direttamente legati alle preoccupazioni che ho espresso, sperando che le loro implicazioni più generali emergano spontaneamente.

Rivolgerò quindi la mia attenzione ai seguenti tipi di domande:

1. Quali proprietà del computer hanno potuto portare l'idea dell'uomo come macchina a un nuovo livello di plausibilità? Chiaramente ci sono state altre macchine che, in vari modi, imitavano l'uomo, per esempio le pale meccaniche a vapore. Ma fino all'invenzione del computer digitale non ci sono state macchine che sapessero svolgere funzioni intellettuali, neanche di portata limitata; vale a dire, macchine che potessero in alcun senso essere definite intelligenti. Ora l'intelligenza artificiale è una sotto-disciplina della scienza dei computer. Questo campo meriterà una discussione particolare più avanti. In ultima analisi, è necessario tracciare una linea di divisione tra intelligenza dell'uomo e intelligenza della macchina. Se questa linea non c'è, allora i sostenitori della psicoterapia computerizzata potrebbero essere semplicemente gli araldi di un'epoca in cui finalmente si riconoscerà che l'uomo non è altro che un complicato meccanismo. In tal caso, sarebbe necessario prevedere e discutere con la massima urgenza le conseguenze di una realtà simile.

2. Il fatto che molte persone si leghino alle macchine con forti relazioni emotive, di per sé non dovrebbe essere sorprendente. Gli strumenti che l'uomo usa diventano, dopo tutto, estensioni del suo corpo. Cosa ancora più importante, per poter usare con abilità i suoi strumenti l'uomo deve interiorizzarne certi aspetti sotto forma di abitudini cinestetiche e percettive. Almeno in quel senso, i suoi strumenti diventano letteralmente parte di lui e lo modificano, alterando così le basi della sua relazione affettiva con se stesso. Ci si può dunque attendere che l'attenzione dell'uomo si concentri più intensamente su strumenti direttamente affini alle sue funzioni intellettuali, cognitive ed emotive, piuttosto che su macchine che semplicemente estendono la sua potenza muscolare. L'intero ambiente dell'uomo occidentale è oggi pervaso da complica-

[6] J. Weizenbaum, 1972.

te protesi tecnologiche per ciascuna sua capacità funzionale. Siccome è un animale dotato di enorme capacità di adattamento, l'uomo è stato capace di accettare queste basi tecnologiche per il rapporto con se stesso, per la sua identità, come autenticamente naturali (cioè, fornite dalla natura). Forse questo aiuta a capire perché egli non si ponga il problema dell'opportunità o meno di confidare le sue sensazioni più intime a un computer.

Ma, allora, una spiegazione di questo tipo suggerirebbe anche che la macchina calcolatrice rappresenti soltanto un'estrapolazione estrema di una molto più generale «usurpazione tecnologica» della capacità dell'uomo di comportarsi come un agente autonomo, nel dare significato al suo mondo. È perciò importante indagare sui sensi più lati in cui l'uomo ha finito per cedere la sua autonomia a un mondo visto come macchina.

3. È forse paradossale che, proprio quando nel senso più profondo l'uomo ha smesso di credere (non parliamo neanche di confidare) nella sua autonomia, egli abbia iniziato ad affidarsi a macchine autonome, vale a dire a macchine che operano per lunghi periodi di tempo interamente sulla base della loro realtà interna. Se il suo affidarsi a tali macchine vuole basarsi su qualcosa che non sia semplicemente inconsolabile disperazione o fede cieca, egli deve spiegare a se stesso che cosa queste macchine fanno, e come fanno a farlo. Questo gli impone di strutturare una qualche concezione delle loro «realtà» interne. Eppure, la maggior parte delle persone non ha la minima idea di che cos'è un computer. Quindi tutte queste persone, a meno che non siano capaci di grande scetticismo (del tipo di quello che viene spontaneo guardando un prestigiatore), potranno spiegare le imprese intellettuali del computer soltanto applicando l'unica analogia che hanno a disposizione, cioè il loro modello della propria capacità di pensare. Non stupisce, quindi, che esagerino nell'apprezzamento; è davvero impossibile immaginare un essere umano capace, per esempio, di imitare *Eliza*, ma per il quale le capacità linguistiche di *Eliza* fossero il limite massimo arrivabile. Di nuovo, la macchina calcolatrice è soltanto un esempio estremo di un fenomeno molto più generale. Anche l'ampiezza di notazione propria dell'uso ordinario della parola «macchina», per quanto grande sia, non basta a dare l'idea della sua vera generalità. Perché oggi, quando parliamo per esempio della burocrazia, o dell'università, o di qualunque altra struttura sociale o politica, l'immagine che generiamo è troppo spesso quella di un processo autonomo simile a quello di una macchina.

Questi, dunque, sono i pensieri e i dubbi che non mi hanno abbandonato da quando mi fu chiaro il significato profondo delle reazioni che ho descritto di fronte a *Eliza*. Eppure, non credo che avrebbero potuto imprimersi così nella mia mente se non perché ero (e sono tuttora) profondamente coinvolto in un concentrato di società tecnologica, in quanto insegnante in quel tempio della tecnologia che è il Massachusetts Institute of Technology (MIT), un'istituzione che si vanta con fierezza di essere «polarizzata attorno alla scienza e alla tecnologia». Qui vivo e lavoro con colleghi, molti dei quali credono che soltanto la scienza moderna permetta una conoscenza attendibile del mondo. Con

loro discuto delle proposte di ricerca da avanzare alle agenzie governative, e in special modo al Dipartimento della «Difesa». A volte sono veramente spaventato quando guardo quello che ci spingiamo a proporre, e la natura delle argomentazioni che escogitiamo per appoggiare le nostre proposte. C'è poi anche il fatto che mi trovo costantemente a confronto con gli studenti, alcuni dei quali hanno già rifiutato tutte le strade non scientifiche per arrivare alla conoscenza del mondo, e vanno soltanto in cerca di un indottrinamento più profondo e dogmatico in questa loro fede (anche se questa parola è stata eliminata dal loro vocabolario). Altri studenti hanno il sospetto che neppure l'intera collezione di macchinari e strumenti del MIT possa aiutarli in modo significativo a dare un senso alla loro vita. Essi avvertono la presenza di un dilemma in un'educazione polarizzata attorno a scienza e tecnologia, un'educazione che pretende implicitamente di aprire un sentiero privilegiato per avvicinarsi alla realtà dei fatti, ma non sa dire loro come riconoscere che cosa va considerato «fatto». Pur riconoscendo l'effettiva importanza di imparare il loro mestiere, essi si ribellano all'idea di lavorare su progetti che non sembrano indirizzati a risolvere problemi pratici interessanti, né a trovare una risposta a problemi teorici.

Questi confronti con la mia realtà sociale quotidiana mi hanno progressivamente convinto che la mia esperienza personale con *Eliza* era sintomatica di problemi più profondi. Sarebbe venuto il momento, ne ero certo, in cui non sarei più stato capace di partecipare a proposte di ricerca o a conferenze, o di rispondere onestamente al bisogno di terapia (sì, è questa la parola giusta) dei miei studenti, senza prima cercare di dare un senso al quadro che la mia esperienza personale con i computer aveva dipinto così bruscamente per me.

Naturalmente, l'introduzione dei computer nella nostra società già altamente tecnologica ha — come cercherò di mostrare — solamente rafforzato e amplificato quelle spinte precedenti che hanno portato l'uomo a una visione sempre più razionalistica della sua società, e a un'immagine sempre più meccanicistica di se stesso. È perciò importante che io strutturi la mia discussione sull'impatto del computer sull'uomo e sulla sua società in modo tale che esso possa essere visto come una speciale codificazione di un impatto molto più esteso, cioè quello sul ruolo dell'uomo di fronte a tecnologie e tecniche che può, a volte, non capire e controllare. Discussioni su questo tema ce ne sono state molte, in passato. E si sono intensificate negli ultimi anni.

Persone di mentalità, carattere, interessi e preparazione molto diversi hanno tutte — pur partendo da posizioni differenti, ed essendo anche in disaccordo su altre questioni di vitale importanza — espresso la loro seria preoccupazione per la situazione creata dalla irrefrenabile marcia di scienza e tecnologia; tra loro ci sono Mumford, Arendt, Ellul, Roszack, Comfort e Boulding. Solo di recente il computer ha cominciato a essere citato in queste discussioni. Ora ci sono segni che indicano lo sviluppo di un dibattito in grande scala sul computer. Per dirla in poche parole, da una parte ci sono coloro che ritengono che i computer possano, debbano e finiranno per fare tutto; dall'altra coloro che, come me, credono che ci siano dei limiti a ciò che dovremmo far fare ai computer.

A prima vista, questo potrebbe sembrare un dibattito «casalingo», con conseguenze modeste e limitate a un piccolo gruppo di specialisti di computer. Ma in ultima analisi, indipendentemente dal modo in cui può venir travestito usando un linguaggio tecnologico gergale, il problema è quello di stabilire se sia o meno possibile ridurre qualunque aspetto del pensiero umano a formalismo logico; o, per usare un'espressione più moderna, se il pensiero umano sia o meno calcolabile. Questo problema ha occupato, sotto diverse forme, pensatori di tutti i tempi. L'uomo ha sempre anelato a principi che potessero organizzare la sua esistenza e darle un significato. Ma, prima che la scienza moderna partorisse le tecnologie che hanno dato realtà e concretezza ai suoi sistemi altrimenti astratti, i sistemi di pensiero che definivano la posizione dell'uomo nell'universo erano sostanzialmente giuridici. Servivano a definire i doveri dell'uomo nei confronti dei suoi simili e della natura. La tradizione giudaica, per esempio, si basa sull'idea di un rapporto contrattuale tra Dio e l'uomo. Questo rapporto deve lasciare, e lascia in effetti, uno spazio di autonomia tanto per Dio quanto per l'uomo, giacché un contratto è un accordo sottoscritto volontariamente da parti che avrebbero la libertà di non farlo. L'autonomia dell'uomo, e i suoi doveri corrispondenti, sono un aspetto centrale di tutti i sistemi religiosi. Le cosmologie spirituali generate dalla scienza moderna, d'altra parte, sono contaminate dal germe della necessità logica. Quando escono dal circolo ristretto degli scienziati e dei filosofi più saggi, esse non si accontentano più di spiegazioni delle apparenze, ma pretendono di saper dire come stanno davvero, e come devono necessariamente stare, le cose. In poche parole, trasformano la verità in dimostrabilità.

Una conseguenza di questo indirizzo della scienza moderna è che la domanda «Quali aspetti della vita sono formalizzabili?» è stata trasformata da domanda morale («Come e in che forma si possono conoscere doveri e responsabilità dell'uomo?») in un interrogativo del tipo «Di quale genere tecnologico l'uomo è una specie?». Anche certi filosofi il cui istinto si ribella all'idea che l'uomo sia completamente comprensibile come una macchina si sono arresi allo spirito dei nostri tempi. Hubert Dreyfus, per esempio, esercita l'artiglieria pesante della fenomenologia sul modello computerizzato dell'uomo [7]. Egli però limita la sua discussione al problema tecnico di cosa i computer possono o non possono fare. Io sosterrei che, anche qualora i computer sapessero imitare l'uomo da tutti i punti di vista — il che in realtà non è vero —, sarebbe opportuno, e anzi urgente, esaminare il computer alla luce dell'eterno bisogno che l'uomo ha di trovare la sua collocazione nel mondo. Il risultato di questa discussione, per quanto riguarda argomenti pratici e di vitale importanza per tutti, dipende dal modo in cui la discussione stessa viene svolta.

Una delle posizioni che intendo sostenere è, in apparenza, ingannevolmente ovvia: è quella secondo cui ci sono differenze importanti tra l'uomo e la macchina visti come entità pensanti. Io sosterrei che, per quanto intelligente possa essere resa una macchina, esistono atti di pensiero che *dovrebbero* essere esclu-

[7] H.L. Dreyfus, *What Computers Can't Do* («Ciò che i computer non possono fare»), Harper and Row, London 1972.

sivo appannaggio degli uomini. Dunque, una questione di rilievo sociale che voglio sollevare riguarda la collocazione da dare al computer nell'ordinamento sociale. Ma, come vedremo, questo problema va al di là del computer, in quanto ha in ultima analisi a che fare con la logicità stessa (indipendentemente dal fatto che la logicità sia codificata o meno in programmi per computer).

Possiamo perdonare al lettore profano, non abituato alla mentalità dell'ambiente scientifico, la sua incredulità di fronte all'idea che ci sia chi sostiene che il pensiero umano sia completamente riducibile a calcoli. Ma proprio questa incredulità può di per sé essere un segno del modo meravigliosamente sottile e seducente in cui la scienza moderna ha finito per influenzare la costruzione immaginaria della realtà da parte dell'uomo.

Certamente, gran parte di ciò che oggi consideriamo benefico e utile, e anche di ciò che chiameremmo conoscenza e saggezza, lo dobbiamo alla scienza. Ma la scienza può anche essere vista come una droga che dà assuefazione. Non soltanto ci siamo nutriti di scienza fino a diventarne dipendenti ma, come succede con molte altre droghe prese in dosi crescenti, la scienza si è gradualmente trasformata in un veleno a effetto ritardato. A partire forse dal fraintendimento della promessa scientifica da parte di Francesco Bacone, l'uomo è stato sedotto fino a sognare — e a impegnarsi per favorire — la creazione di un'era della razionalità, ma con una visione così contorta della razionalità da fargliela sembrare equivalente alla logicità. Così siamo arrivati quasi al punto in cui qualunque vero dilemma umano viene visto come un semplice paradosso, come una contraddizione soltanto apparente, che potrebbe essere risolta se si applicasse giudiziosamente la fredda logica propria di una razionalità superiore. Addirittura, si è arrivati a pensare che certe guerre sanguinose fossero soltanto problemi da risolvere usando schiere di «risolutori di problemi» professionisti. Come ha detto Hannah Arendt a proposito dei creatori e degli esecutori della politica del Pentagono in anni recenti:

> Non erano soltanto intelligenti, ma si vantavano di essere « razionali » [...] Erano ansiosi di trovare formule, preferibilmente espresse in un linguaggio pseudomatematico, capaci di unificare i fenomeni più disparati con cui la realtà si presentava loro; vale a dire, desideravano scoprire delle *leggi* con cui spiegare e prevedere fatti storici e politici, come se questi fossero necessari, e quindi affidabili, tanto quanto i fenomeni naturali lo sono dal punto di vista dei fisici [...] [Essi] non *giudicavano*; calcolavano [...] Una fiducia completamente irrazionale nella calcolabilità della realtà [divenne] il *leit-motiv* per il processo decisionale [8].

Allo stesso modo, quasi tutti i contrasti politici, come quelli tra razze o tra una popolazione e il suo governo, hanno finito per essere considerati come meri fallimenti comunicativi. Simili crepe nel tessuto sociale possono allora essere sistematicamente riparate, applicando sapientemente le più recenti tecniche di elaborazione delle informazioni; almeno, così si crede. E così l'equazione razionalità = logicità — che la scienza proprio in virtù del suo successo ci

[8] H. Arendt, *Crises of the Republic* («Crisi della Repubblica»), Harcourt Brace Jovanovich, Harvest edition 1972, pp. 11 ss.

ha persuaso ad adottare come se fosse un assioma — ci ha portato a negare l'esistenza del conflitto umano e della possibilità di far convergere interessi umani incommensurabili e valori umani disparati; in fondo, ci ha portato a negare l'esistenza stessa di valori umani.

È possibile che i valori umani siano illusori, come peraltro sostiene B.F. Skinner. Se lo sono, allora presumibilmente sta alla scienza dimostrarlo, e infatti Skinner (in quanto scienziato) cerca di farlo. Ma allora, la scienza dev'essere anch'essa un sistema illusorio. Infatti, l'unica conoscenza sicura che la scienza ci può dare è la conoscenza del comportamento di sistemi formali, cioè di sistemi che sono giochi inventati dall'uomo; e in questi, affermare la verità non è diverso dall'affermare che, come in una partita di scacchi, una particolare configurazione di pezzi sulla scacchiera è stata raggiunta attraverso una sequenza di mosse lecite. Quando la scienza pretende di fare affermazioni sulle esperienze dell'uomo, le basa su identità tra gli oggetti primitivi (cioè indefiniti) di uno dei suoi formalismi, i pezzi di uno dei suoi giochi e qualche insieme di osservazioni umane. Non sarà mai possibile dimostrare che un simile insieme di corrispondenze sia corretto. Al massimo, sarà possibile falsificarlo, nel senso che manipolazioni formali dei simboli di un sistema possono portare a configurazioni simboliche che, quando siano lette alla luce dell'insieme di corrispondenze in questione, producono interpretazioni contrarie a fenomeni osservati empiricamente. Dunque, tutta la scienza empirica è un'elaborata struttura costruita su pali ancorati non già su un fondale roccioso, come comunemente si crede, bensì sulle sabbie mobili del fallibile giudizio umano, della congettura e dell'intuizione. Non è neppure vero, di nuovo contrariamente alla credenza comune, che un singolo presunto controesempio (che, se accettato come vero, falsificherebbe sicuramente una certa teoria scientifica) porti in generale all'immediato abbandono di quella teoria. Probabilmente, tutte le teorie scientifiche attualmente accettate dagli scienziati stessi (tranne quelle teorie puramente formali che non sostengono di avere alcun rapporto col mondo empirico) si trovano oggi di fronte ad evidenze contraddittorie di peso più che trascurabile, le quali — se si basasse il giudizio su di esse — avrebbero il logico effetto di invalidare tali teorie. Evidenze di quel tipo sono spesso spiegate (vale a dire, messe da parte) attribuendole a un errore di qualche tipo, per esempio nell'osservazione, o caratterizzandole come inessenziali, o in base all'ipotesi (vale a dire, all'atto di fede) che un giorno si scoprirà che quei fatti, apparentemente in contraddizione con la teoria scientifica, possono invece essere incorporati nella teoria pur conservando la loro peculiarità. In questo modo, gli scienziati continuano ad affidarsi a teorie già messe in dubbio, e a dedurne «fatti scientifici» [9].

[9] Così, Charles Everett scrive della teoria della combustione basata sul flogisto, e oggi scartata (nell'*Encyclopaedia Britannica*, 11ma ed., 1911, vol. VI, p. 34): «Le obiezioni degli scienziati contrari all'idea del flogisto, come il fatto che le calci pesano più dei metalli originali, anziché meno come suggerito dalla teoria, venivano ribattute postulando che il flogisto fosse un elemento di leggerezza, oppure erano ignorate come incidenti di percorso, in quanto il cambiamento di qualità era considerato l'unico aspetto importante della questione». Everett cita H. Cavendish e J. Priestley, entrambi grandi scienziati del loro tempo, come sostenitori della teoria del flogisto.

L'uomo della strada certamente crede che questi fatti scientifici siano tanto ben assodati e dimostrati quanto la sua stessa esistenza. La sua certezza è un'illusione. E neppure lo scienziato è immune dalla stessa illusione. All'atto pratico egli deve, dopo tutto, dimenticare molti dubbi, se vuole fare o pensare qualcosa. È un po' come uno spettatore a teatro il quale, per partecipare a ciò che succede sulla scena e capirlo, deve almeno per qualche istante fare finta di trovarsi di fronte a eventi reali. Lo scienziato deve credere nella sua ipotesi di lavoro, con la vasta struttura di teorie e ipotesi che essa sottende, se non altro per poter affrontare la discussione. Spesso questa «discussione» lo tiene impegnato per la vita intera. Progressivamente, egli diventa ciò che all'inizio faceva solo finta di essere: un vero credente. Ho scelto deliberatamente la parola «discussione», perché le dimostrazioni scientifiche, e anche le prove matematiche, sono fondamentalmente degli atti di persuasione.

Le affermazioni scientifiche non possono mai essere certe; possono soltanto essere più o meno credibili. E la credibilità riguarda la psicologia individuale, ha cioè senso se riferita a un osservatore individuale. Dire che una proposizione è credibile significa, dopo tutto, dire che è creduta da un agente libero di non crederla, cioè da un osservatore il quale, dopo aver esercitato il suo giudizio e (eventualmente) la sua intuizione, sceglie di accettare la proposizione come degna di essere creduta. Come può quindi la scienza, che si basa essa stessa sicuramente, in fin dei conti, sull'insieme di molti giudizi di valore umani, dimostrare che i giudizi di valore umani sono illusori? Non può farlo, a meno di rinunciare alla sua stessa prerogativa di unica strada legittima per la comprensione dell'uomo e del suo mondo.

Ma nessuna argomentazione puramente logica, per quanto stringente e convincente possa essere, può modificare la realtà, e cioè il fatto che, secondo il buon senso comune, la scienza è diventata l'unica forma legittima di conoscenza. Quando dico che la scienza è stata progressivamente trasformata in un veleno a effetto ritardato, intendo dire che la patente di certezza data alla conoscenza scientifica dal senso comune, in modo così universale da costituire ormai un dogma, ha sostanzialmente delegittimato ogni altro modo per comprendere. Una volta le arti, e in particolare la letteratura, erano viste come fonti di nutrimento e di comprensione intellettuale, mentre oggi sono per lo più considerate come forme di divertimento. Il teatro greco e orientale antico, le opere di Shakespeare, le scene popolate dai vari Ibsen e Cechov in tempi più vicini a noi: queste erano scuole. I programmi che spiegavano erano dei veicoli per la comprensione della società che rappresentavano. Oggi, anche se occasionalmente qualche Arthur Miller o Edward Albee sopravvive e viene autorizzato a insegnare sulle scene di New York o di Londra, la gente è affamata soltanto di ciò che le viene presentato come conoscenza dimostrata scientificamente. Cercano di saziarsi nei supermercati scientifici come la rivista «Psychology Today», o nelle versioni divulgative dei lavori di Masters e Johnson, o nella scientologia rivelata da L.Ron Hubbard. La fede nell'equazione razionalità = logicità ha corroso il potere profetico del linguaggio stesso. Sappiamo contare, ma ci dimentichiamo rapidamente di come si faccia a riconoscere cosa valga la pena di contare, e perché.

1
SUGLI STRUMENTI

Le vicende dell'uomo e quelle delle sue macchine sono collegate inseparabilmente fra loro. Le macchine hanno messo l'uomo in grado di trasformare il suo ambiente fisico. Con il loro aiuto egli ha arato la terra, costruito città e scavato grandi canali. Simili trasformazioni dell'habitat dell'uomo hanno necessariamente determinato cambiamenti della sua organizzazione sociale. Ma, forse in modo ancora più cruciale, le macchine dell'uomo hanno fortemente influenzato la sua visione del mondo e, quindi, di se stesso. L'uomo è consapevole di se stesso, dell'esistenza di altri esseri simili a lui e di un mondo che è — almeno in una certa misura — malleabile. Cosa più importante di tutte, l'uomo sa prevedere. Nel progettare strumenti per dissodare il terreno, egli ne mette alla prova la funzionalità con la sua immaginazione. In più, sapendo di essere una creatura sociale e di dovere inevitabilmente morire, egli è per necessità un insegnante. I suoi utensili, qualunque ne sia la primaria funzione pratica, sono per necessità anche strumenti pedagogici. Sono, quindi, parte del materiale dal quale l'uomo estrae la propria immaginaria ricostruzione del mondo. È all'interno del mondo sociale e intellettuale che egli stesso crea che l'individuo prova e riprova infinite volte la messa in scena di come il mondo avrebbe potuto essere, e di cosa potrebbe diventare. Quel mondo è il ricettacolo della sua soggettività. Perciò, è stimolo per la sua consapevolezza e, infine, forza costruttrice dello stesso mondo materiale. È questo mondo auto-costruito che l'uomo incontra come una forza in apparenza esterna. Ma in realtà lo ha dentro di sé; ciò che ha di fronte è il suo personale modello di universo e, dal momento che egli ne è parte, il suo modello di se stesso.

Ci sono poche cose che l'uomo può creare senza prima immaginare di crearle. Pensando all'uomo dell'età della pietra che faticosamente si costruisce l'ascia, possiamo immaginare come, dentro di sé, egli «provi» l'uso che ne farà. Non abbiamo forse ripetuto tutti questa esperienza ancestrale quando, da bambini, costruivamo giocattoli primitivi con qualunque materiale avessimo a portata di mano? Ma utensili e macchine non testimoniano solamente l'immaginazione umana con la sua portata creativa, e certamente non sono importanti soltanto come strumenti per la trasformazione di una terra malleabile: sono simboli pregnanti di per sé. Simbolizzano le attività che rendono pos-

sibili, vale a dire l'uso che se ne può fare. Un remo è uno strumento per remare, e rappresenta la capacità di remare in tutta la sua complessità. Nessuna persona che non abbia mai remato può vedere un remo per quello che realmente è. Così come chi non l'abbia mai suonato vedrà il violino in modo completamente diverso da un violinista. Un utensile è anche un copione per nuove messe in scena delle azioni che simbolizza. In questo senso è uno strumento pedagogico, un veicolo per informare uomini di altri tempi e luoghi di certi modi di pensare e agire acquisiti per processo culturale. L'utensile come simbolo, da questo punto di vista, trascende così il suo ruolo di mezzo pratico per certi fini: è invece un componente della ricreazione simbolica che l'uomo fa del suo mondo. Deve perciò inevitabilmente entrare a far parte del fantasioso progetto che in continuazione costruisce il suo mondo. In tal senso, dunque, l'utensile è molto più di un semplice apparecchio: è un agente di cambiamento. Ed è più che un frammento di un disegno di mondo concepito per l'uomo e trasmessogli dai suoi antenati. Benché sia anche quello.

È facile rendersi conto che gli utensili manuali, e specialmente le armi manuali, hanno effetti diretti sulla fantasia degli individui che li usano. Quando i cacciatori sono venuti in possesso della lancia, per esempio, devono essersi visti in un rapporto completamente nuovo con il loro mondo. Grandi animali che fino allora avevano razziato i loro depositi di cibo e attaccato i loro bambini, e di cui avevano terrore, diventavano ora una preda. La fonte di cibo per l'uomo crebbe, perché ora era possibile uccidere animali a distanza, comprese molte specie che avevano evitato il contatto con l'uomo. La maggiore abbondanza di cibo deve anche aver allargato la porzione di mondo in cui muoversi, aumentando così la probabilità di incontrare altre razze umane. Cambiò allora l'esperienza che l'uomo aveva del mondo, e dovette cambiare anche l'idea di quale fosse il suo ruolo in questo mondo.

La carabina a sei colpi del West americano nel diciannovesimo secolo era chiamata «grande livella», nome che testimonia in modo eloquente l'effetto che quel pezzo di metallo ebbe sull'immagine di sé di quei «pistoleros» che, se spogliati delle loro armi, si sentivano svantaggiati rispetto ai loro concittadini. Ma apparecchi e macchinari, magari conosciuti (e certamente posseduti e fatti funzionare) soltanto da un numero relativamente esiguo di membri della società, hanno spesso influenzato l'immagine di sé degli individui, e l'immagine del mondo della società nel suo complesso, altrettanto profondamente quanto gli utensili manuali di vasta diffusione. Tutti i diversi tipi di nave, per esempio, sono stati un mezzo per rendere l'uomo consapevole della vastità del suo territorio. Hanno permesso a culture diverse di incontrarsi e di fecondarsi a vicenda. Le imbarcazioni dei marinai e gli altri loro artefatti, i loro miti e le loro leggende ne hanno trasmesso il patrimonio tradizionale efficacemente di generazione in generazione. E hanno influenzato il subconscio di chi stava a terra, non meno di quello di chi davvero navigava. La carta stampata ha trasformato il mondo anche per quei milioni di persone che, ad esempio all'epoca di Martin Lutero, rimasero analfabeti e non videro magari un libro in tutta la loro vita. E delle grandi masse di persone la cui vita venne direttamente e completamente trasformata dalla rivoluzione industriale, quante vennero mai

davvero a contatto con una macchina a vapore? E neanche la società moderna è immune da enormi contraccolpi ricevuti come effetti collaterali dell'introduzione di nuove macchine. La macchina per raccogliere il cotone fu introdotta nelle piantagioni del Sud degli Stati Uniti a cominciare dal 1955. Rapidamente distrusse il mercato per l'unica merce che grandi masse di braccianti neri del Sud avevano da vendere: il loro lavoro. Incominciò così l'emigrazione di massa dei neri americani verso le città, in particolare verso centri industriali del Nord come Detroit, Chicago e New York, ma anche verso le grandi città del Sud, come Birmingham e Atlanta. Sicuramente questo enorme cambiamento nel panorama demografico degli Stati Uniti, questa migrazione interna di milioni di cittadini, è stato e rimane uno dei principali elementi a determinare il corso del movimento per i diritti civili negli Stati Uniti. E quel movimento ha influenzato in modo non trascurabile la consapevolezza per lo meno di ogni cittadino americano, se non di ogni essere umano adulto sulla Terra.

In che cosa consiste quest'implacabile impulso della macchina, che la fa insinuare al centro del materiale con cui l'uomo costruisce il suo mondo?

Molte macchine sono delle aggiunte funzionali al corpo umano, sostanzialmente delle protesi. Alcune, come la leva e le macchine a vapore, aumentano la potenza muscolare grezza di coloro che le fanno funzionare; alcune, come il microscopio, il telescopio e vari strumenti di misura, sono estensioni dell'apparato sensoriale umano. Altre estendono la portata fisica dell'uomo. La lancia e la radio, per esempio, permettono all'uomo di proiettare la propria influenza su una distanza che supera quella cui arrivano le sue braccia e la sua voce. I veicoli dell'uomo gli permettono di viaggiare più velocemente e più lontano di quanto potrebbe fare con le sue gambe, e gli permettono di trasportare grandi pesi su lunghe distanze. È facile riconoscere come e perché simili appendici meccaniche rafforzino direttamente il senso di potere che l'uomo ha sul mondo materiale. Ed esse hanno anche un importante effetto psicologico: dicono all'uomo che egli può scegliere di rifarsi diverso. Infatti, esse fanno parte di quell'insieme di simboli che l'uomo usa per ricreare il suo passato, cioè per costruire la sua storia, e per creare il suo futuro. Testimoniano che l'uomo, il costruttore, può andare al di là dei limiti che gli sono imposti dalla debolezza del suo corpo e dei suoi sensi. Una volta, l'uomo poteva uccidere un altro animale soltanto colpendolo o afferrandolo con le proprie mani; poi si è costruito l'ascia, la lancia, la freccia, la pallottola sparata da una pistola, la bomba. Ora, cariche esplosive montate su missili possono distruggere l'umanità stessa. E questo dà la misura di come l'uomo si sia esteso e trasformato, da quando incominciò a costruire utensili.

Il ricondurre l'influenza delle protesi meccaniche sulla trasformazione dell'uomo unicamente alla potenza che gli hanno permesso di accumulare può spingere a una visione del rapporto uomo-natura la cui principale — anzi, quasi unica — componente è un grezzo rapporto di forze. Da questo punto di vista, l'uomo ha infine conquistato la natura semplicemente dotandosi di un potere sufficiente a fargli vincere spazio e tempo naturali, a progettare vita e morte, a distruggere infine del tutto la Terra. Ma questa idea è sbagliata, anche se accettiamo che il sogno eterno dell'uomo sia stato non solo la scoperta della

natura, ma la sua conquista, e che questo sogno sia oggi in larga misura realizzato. Perché se è vero che la vittoria sulla natura è stata conseguita in questa nostra epoca, è altresì vero che la natura su cui l'uomo oggi regna è molto diversa da quella in cui l'uomo viveva prima della rivoluzione scientifica. Anzi, il trucco che l'uomo ha usato, e che ha permesso il sorgere della scienza moderna, non è stato altro che la trasformazione della natura e della percezione che l'uomo stesso ha della realtà.

Il cambiamento straordinario nella vita mentale dell'uomo, a partire circa dal quattordicesimo secolo, è stato nella sua percezione del tempo e, di conseguenza, dello spazio. L'uomo aveva osservato molto tempo fa (e possiamo supporre che ci avesse riflettuto sopra) delle regolarità nel mondo che lo circondava. Alexander Marshack ha dimostrato che anche l'uomo del Paleolitico superiore (circa 30.000 a.C.) aveva una notazione per indicare i cicli lunari che era, per usare le parole di Marshack, «già evoluta, complessa e sofisticata, una tradizione che sembrerebbe essere già vecchia di migliaia di anni in quell'epoca» [1]. Ma dall'antichità classica fino a tempi relativamente recenti, la regolarità dell'universo venne ricercata e percepita in forma di grandi armonie tematiche. L'idea che la natura si comporti sistematicamente nel modo in cui l'intendiamo noi — vale a dire, che ogni parte e ogni aspetto della natura possa essere isolato come sottosistema governato da leggi descrivibili in funzione del tempo —, questa idea non avrebbe neppure potuto essere compresa da uomini che percepivano il tempo non già come una sequenza di unità astratte (cioè ore, minuti e secondi), ma come una sequenza di eventi che ricorrono costantemente.

Le ore del giorno venivano identificate con eventi come il passaggio del sole sulla verticale di una particolare colonna di pietre o, come ci dice Omero, con l'inizio e la fine di particolari lavori, come il mettere e il togliere il giogo al bue (mattino e sera, rispettivamente). Gli intervalli di tempo venivano indicati riferendosi a occupazioni comuni, per esempio il tempo necessario per arrivare a piedi in un luogo noto, o per far bollire una data quantità di acqua. Le stagioni venivano riconosciute attraverso avvenimenti stagionali ricorrenti, per esempio la partenza degli uccelli. Fino a quando la teoria dell'evoluzione di Darwin non incominciò a far parte del bagaglio di idee generalmente accettate, cioè a diventare «buon senso comune», la gente sapeva che il mondo che la circondava — con piante e animali che si riproducono, con fiumi che scorrono, seccano e scorrono di nuovo, con mari che pulsano in grandi ritmi di marea, con l'ininterrotto spettacolo del cielo — era sempre esistito, e che la sua legge fondamentale era l'eterna periodicità. Il tempo cosmologico, così come il tempo percepito nella vita quotidiana, era perciò una sorta di complicato pulsare, un ripetersi ed echeggiare di eventi. Possiamo forse capire questo modo di pensare ascoltando, per fare un esempio, le grandi fughe di Bach. Ma ci è richiesta qui una particolare forma di ascolto: non dobbiamo pensare in modo moderno, vale a dire a Bach come «risolutore di problemi», o a ciascu-

[1] A. Marschak, *The Roots of Civilization* («Le radici della civiltà»), Macmillan, New York 1972, p. 57.

na delle sue opere nell'*Arte della fuga* come alla «soluzione», il più possibile raffinata, di un problema che egli si fosse inizialmente posto. Dobbiamo invece pensare che Bach avesse in mente l'intero piano dell'opera già durante il lavoro, che pensasse l'*Arte della fuga* come un'opera unitaria, senza un inizio né una fine, eterna come lo è il cosmo, e come questo enormemente intricata nelle sue connessioni, cerchi interni a cerchi interni ad altri cerchi. Allora potremmo forse pensare a questa musica come a qualcosa che non ha semplicemente scandito la vita, ma l'ha permeata interamente, tanto sulla grande scala teologico-cosmologica quanto sul modesto piano della quotidianità. Il tempo che essa rappresenta è una rivoluzione di cicli, e di epicicli all'interno di ogni ciclo, ben diversa dal ricettacolo di una progressione uniforme di momenti astratti, che noi oggi lo «sappiamo» essere. E la natura stessa consisteva certamente di fenomeni individuali, che erano però metamorfosi di se stessi in costante ripetizione, ed erano perciò permanenti, eterni. «Ciò che è eterno è circolare, e ciò che è circolare è eterno» disse Aristotele; e anche Galileo credeva ancora che l'universo fosse eterno e governato da ricorsi e periodicità.

La concezione darwiniana del tempo era radicalmente diversa. Egli vedeva la natura stessa come un processo in atto nel tempo, e i fenomeni individuali della natura come metamorfosi irreversibili. Ma egli era tutt'altro che l'iniziatore dell'idea di progresso che è oggi così presente in noi. Anzi, non avrebbe potuto pensare ciò che pensava, se qualcosa di molto simile alla nostra attuale idea di tempo non fosse stato già allora parte integrante del buon senso comune.

Il modo in cui la percezione del tempo è cambiata dai tempi dei nostri antenati a oggi getta luce sul ruolo svolto da un altro tipo di macchina (non «protetica») nella trasformazione dell'uomo da creatura della, e vivente nella, natura a signore della natura.

L'orologio non è una macchina protetica; ciò che produce non è un'estensione dei muscoli o dei sensi umani, bensì ore, minuti, secondi, oggi persino micro-, nano- e picosecondi. Secondo Lewis Mumford è l'orologio, non il motore a vapore, «la macchina-chiave dell'era industriale moderna» [2]. Nel brillante capitolo iniziale del suo *Tecnica e civiltà* egli descrive, tra l'altro, il modo in cui nel Medio Evo la vita ordinata dei monasteri influenzava la vita delle comunità loro adiacenti.

> Il monastero era sede di una vita regolare... l'abitudine dell'ordine stesso e la seria regolazione delle sequenze temporali era diventata quasi una seconda natura per il monastero [...]. I monasteri — a un certo punto ve ne furono 40.000 sotto la regola benedettina — contribuivano a dare all'impresa umana il ritmo e il battito regolare collettivo proprio della macchina; perché l'orologio non è solo un modo per tener conto delle ore, ma anche per sincronizzare le azioni degli uomini [...] nel tredicesimo secolo ci sono precise descrizioni di orologi meccanici, e prima del 1370 un orologio «moderno» ben progettato era stato co-

[2] L. Mumford, *Technics and Civilization* («Tecnica e civiltà»), Harcourt Brace Jovanovich, New York 1963, p. 14.

struito a Parigi da Heinrich von Wyck. Nel frattempo, erano venuti alla luce i primi campanili, e i nuovi orologi, sebbene non avessero [fino al quattordicesimo secolo] lancette che traducessero il moto del tempo in un moto nello spazio, pur sempre battevano le ore. Le nuvole che potevano paralizzare la meridiana [...] non erano più un ostacolo per il computo del tempo: inverno o estate, giorno o notte, si sentiva ormai il misurato tic-tac dell'orologio. Lo strumento si diffuse rapidamente al di fuori del monastero; e il regolare rintocco delle campane portò una nuova regolarità nella vita dell'artigiano e del mercante. Le campane della torre dell'orologio quasi diventarono definizione dell'esistenza urbana. L'occupazione di chi teneva il computo del tempo si trasformò in quella di controllore, organizzatore, partitore del tempo altrui. E mentre questo avveniva, gradualmente l'Eternità smise di fungere da misura e obiettivo delle azioni umane [3].

Mumford prosegue con l'osservazione cruciale che l'orologio «ha separato il tempo dagli eventi umani e ha contribuito a creare l'idea di un mondo indipendente formato da sequenze misurabili matematicamente: lo speciale mondo della scienza» [4]. È difficile valutare appieno l'importanza di questo effetto dell'orologio sulla percezione umana del mondo. La nostra attuale visione del tempo è così profondamente parte di noi, quasi una nostra «seconda natura», che non siamo sostanzialmente più capaci di identificare il ruolo che essa svolge nel nostro modo di pensare. Alexander Marshack osserva:

> Il concetto di processo soggetto al fattore-tempo nelle scienze pure è oggi quasi tautologico, dacché tutti i processi — semplici o complessi, sequenziali o intercorrelati, finiti o infiniti — si sviluppano o continuano e hanno velocità, durata, periodicità ecc. che sono misurabili o stimabili. Tuttavia, le scienze che studiano questi processi risentono anch'esse del «fattore-tempo», dal momento che i processi di conoscimento e riconoscimento, programmazione, ricerca, analisi, confronto e interpretazione sono anch'essi sequenziali, intercorrelati, sviluppantisi e cumulativi [5].

E infatti le due fondamentali equazioni della fisica che ogni studente conosce sono $F = ma$ e $E = mc^2$. La a nella prima sta per accelerazione, cioè cambiamento di velocità nel tempo, e la c nella seconda rappresenta la velocità della luce, cioè lo spostamento della luce al passar del tempo.

Parlo qui dell'orologio non soltanto perché è stato un elemento cruciale nel determinare il modo di pensare dell'uomo — ci sono state, dopo tutto, molte altre invenzioni che hanno contribuito a dare inizio al nuovo razionalismo scientifico; per esempio, i meridiani e i paralleli sul globo —, bensì per mostrare che non sono state solo le macchine protetiche a dare all'uomo potere sulla natura. L'orologio chiaramente non è una macchina protetica; non amplifica i muscoli né i sensi umani. È una macchina autonoma.

Molte macchine sono automatiche nel senso che, una volta accese, possono funzionare a lungo da sole. Ma, per la maggior parte, devono essere predi-

[3] *Ivi*, pp. 13, 14.
[4] *Ivi*, p. 15.
[5] Marschak, *op. cit.*, p. 14.

sposte per il compito cui sono destinate, e poi controllate e regolate, da sensori o dalla guida umana. Una macchina è automatica se, una volta fatta partire, procede da sola sulla base di un modello di qualche aspetto del mondo reale che contiene al suo interno. Gli orologi sono fondamentalmente dei modelli del sistema planetario, sono la prima macchina autonoma costruita dall'uomo, e fino all'avvento del computer rimasero l'unica veramente importante.

Quando ha iniziato a usare l'orologio per calcolare il tempo, l'uomo non ha più regolato la sua vita quotidiana basandosi esclusivamente, per esempio, sulla posizione del sole rispetto a certe pietre o sul canto del gallo, ma si è invece affidato ad un modello autonomo di un fenomeno naturale. I vari possibili stati di questo modello vennero chiamati con dei nomi, e in questo modo furono materializzati. E il loro insieme si sovrappose al mondo esistente e lo cambiò, proprio come avrebbe potuto fare un cataclisma geografico o climatico. L'uomo doveva ora sviluppare nuovi sensi per trovare la sua strada nel mondo. L'orologio aveva letteralmente creato una nuova realtà; e questo è ciò che intendevo quando ho detto che il trucco usato dall'uomo per spianare la strada all'ascesa della scienza moderna è stato niente di meno che la trasformazione della natura e della sua propria percezione della realtà. È importante rendersi conto che questa realtà di nuova creazione era e rimane una versione impoverita della precedente, giacché si fonda sul rifiuto di quelle esperienze dirette che erano alla base, anzi costituivano la vecchia realtà. Il senso di appetito venne rifiutato come stimolo per mangiare; invece, si mangiava quando un modello astratto aveva raggiunto un certo stato, cioè quando le lancette di un orologio indicavano certi simboli sul suo quadrante (anche qui l'antropomorfismo è assai significativo)[6], e lo stesso accadeva per il sonno e la veglia, e così via.

Questo rifiuto dell'esperienza diretta doveva diventare una delle principali caratteristiche della scienza moderna. Fu inculcato nella cultura europea occidentale non soltanto dall'orologio, ma anche dai molti strumenti sensori protetici, particolarmente da quelli che informavano sui fenomeni che dovevano controllare per mezzo di indicatori la cui posizione veniva alla fine tradotta in numeri. Gradualmente all'inizio, poi sempre più rapidamente e, va detto, quasi obbligatoriamente, le esperienze della realtà dovettero essere rappresentabili con numeri, se volevano apparire legittime agli occhi del buon senso comune. Oggi si crede che la manipolazione enormemente complicata di grandi insiemi di numeri sia in grado di generare nuovi aspetti della realtà. Questi ultimi sono confermati confrontando i numeri appena ottenuti con la lettura delle indicazioni di altri strumenti che mediano fra l'uomo e la natura e che, naturalmente, producono molti altri numeri.

«L'uomo scientifico deve, più di ogni altra cosa, sforzarsi di eliminare le sue opinioni quando giudica» scrisse Karl Pearson nel 1982[7]. Dei molti

[6] L'Autore si riferisce qui al fatto che, in inglese, «lancetta» si dice *hand* («mano») e «quadrante» si dice *face* («faccia»), NdT.

[7] C. Pearson, *The Grammar of Science* («La grammatica della scienza»), Dent, London 1911, p. 11.

scienziati che conosco, pochissimi non sarebbero d'accordo. Eppure, bisogna riconoscere che tale affermazione spinge l'uomo a diventare un'intelligenza impersonale, a diventare egli stesso uno strumento di misura, una macchina. A tanto ha portato l'alleanza, all'inizio così innocente, con protesi e lancette indicatrici. E su una cultura di questo tipo è esplosa l'era del computer.

«Ogni uomo che pensa — ha scritto John Dewey — mette in pericolo una qualche parte di un mondo apparentemente stabile, e nessuno può prevedere che cosa emergerà al suo posto». Lo stesso vale per chiunque inventi un nuovo strumento o — che è lo stesso — trovi una nuova utilizzazione per uno strumento già noto. La lunga prospettiva storica che ci aiuta a comprendere l'antichità classica, il Medio Evo e l'inizio dell'era moderna ci aiuta anche a formulare ipotesi plausibili per spiegare le nuove realtà che, in quei tempi, emersero a sostituire realtà precedenti, messe in pericolo dall'introduzione di nuovi strumenti. Ma quando ci avviciniamo al problema di capire il significato e la portata delle vicende che riguardano, da un lato, il cambiamento nella consapevolezza dell'uomo moderno, dall'altro, lo sviluppo degli strumenti contemporanei, e in particolare del computer, allora la nostra prospettiva necessariamente si appiattisce. Ben poco possiamo fare che non sia proiettare le lezioni derivate dalla nostra comprensione del passato, le nostre ipotesi plausibili, sul presente e sul futuro. E la difficoltà del nostro compito è assai accresciuta dal fatto che gli strumenti moderni hanno sulla società un impatto molto più forte e molto più veloce di quelli del passato.

La spinta data dall'orologio verso l'alienazione dell'uomo dalla natura ha richiesto secoli per penetrare e condizionare profondamente l'umanità nel suo complesso. E anche allora, ha dovuto combinarsi sinergeticamente con molti altri fattori emergenti per poter esercitare la sua influenza. La macchina a vapore è arrivata quando, nella visione del buon senso comune, spazio e tempo erano già quantificati. Una natura eterna governata dalle leggi immutabili della periodicità comportava un mandato d'autorità, reso esplicito nei libri sacri ed esercitato da vicari istituzionali dell'ordine eterno. Tale autorità quasi costituzionale, e quindi vincolata, era stata già da lungo tempo soppiantata, per esempio dalla relativamente svincolata autorità del denaro, vale a dire del valore quantificato, e in particolare del valore del lavoro umano quantificato. Queste e molte altre circostanze si combinarono fino a rendere possibile che la macchina a vapore finisse per trasformare radicalmente la società. Strumenti inventati successivamente, ad esempio telefono, automobile, radio, hanno avuto impatto su una società già resa schiava da quello che gli economisti chiamano «principio del maiale»: se qualcosa è buono, averne di più è meglio. La fame di velocità e di capacità di comunicazione, spesso stimolata dalle macchine stesse, insieme a nuove tecniche di *marketing* ad esse associate, ha reso possibile la loro rapida diffusione nella società e la trasformazione sempre più rapida della società sotto la loro influenza.

Quando venne installata la prima linea telegrafica che collegava il Texas con New York, ci fu chi espresse il dubbio che la gente di quei due posti non avesse molto da dirsi, in fondo. Ma quando il computer digitale emerse dai laboratori universitari ed entrò nelle istituzioni commerciali, militari e indu-

striali americane, non c'erano dubbi sulla sua potenziale utilità. Al contrario, i manager e i tecnocrati americani riconobbero che il computer era arrivato appena in tempo per evitare crisi catastrofiche: non fosse stato per il computer, si disse, sarebbe stato impossibile trovare personale a sufficienza per le banche, i problemi logistici e di comunicazione, sempre più complessi per le forze armate americane sparse su tutta la Terra, non avrebbero potuto essere risolti, e i programmi commerciali e il trasporto di merci non avrebbero potuto essere mantenuti in funzione. La finanza americana aveva di fronte un problema di «comando e controllo» simile a quello incontrato dalla sua controparte militare. E, come il Pentagono, anch'essa era sempre più diversificata e internazionalizzata. La società americana alla fine della Seconda Guerra Mondiale aveva davanti a sé il compito di eseguire calcoli così estesi e complicati da non avere precedenti, e il computer, quasi miracolosamente si direbbe, arrivò giusto in tempo per eseguirli.

In realtà, enormi problemi manageriali, tecnologici e scientifici erano stati risolti senza l'aiuto dei calcolatori elettronici nei decenni precedenti la Seconda Guerra Mondiale, e in particolare durante la guerra stessa. Una frazione molto grande della struttura industriale degli Stati Uniti era coordinata per fornire i beni di prima necessità — cibo, vestiti ecc. — e i mezzi di trasporto necessari a grandi eserciti sparsi su tutto il globo. Il Progetto Manhattan produsse la bomba atomica senza usare calcolatori elettronici; eppure, i problemi scientifici e ingegneristici risolti nel suo ambito richiesero probabilmente più calcoli di quanti ne fossero stati necessari per tutti gli studi astronomici effettuati fino allora. Le dimensioni dell'impresa manageriale che venne condotta sicuramente non furono inferiori a quelle del Progetto Apollo degli anni Sessanta. Probabilmente, quasi tutti oggi credono che il Progetto Apollo non avrebbe potuto essere realizzato senza i computer. La storia del Progetto Manhattan sembra contraddire tale opinione. E ci sono credenze simili per quanto riguarda il bisogno di computer nell'organizzazione di grandi società commerciali e del sistema militare, l'indispensabilità dei computer nei moderni calcoli scientifici e, anzi, l'impossibilità stessa di far proseguire il cammino della scienza e del commercio moderni senza l'aiuto dei computer [8].

La credenza che i computer siano indispensabili non è del tutto infondata. Il computer diventa componente indispensabile di una struttura quando è così diffusamente integrato in essa, così inserito in diverse sotto-strutture vitali, che non lo si può più eliminare senza sbilanciare fatalmente l'intera struttura. Quella appena enunciata è, in sostanza, una tautologia. La sua utilità sta

[8] Sono sicuro che, se computer sofisticati come quelli attuali fossero stati disponibili già nel 1940, i tecnici del Progetto Manhattan avrebbero giurato che anche quell'impresa sarebbe stata impossibile senza computer. E simili convinte testimonianze le avremmo avute dai progettisti degli aerei della Seconda Guerra Mondiale, e dai responsabili degli aspetti logistici di quella guerra. Se la Germania avesse avuto computer fin dall'inizio della dittatura di Hitler, tutti oggi penserebbero che solo grazie al computer i nazisti abbiano potuto controllare il popolo tedesco, e realizzare il trasporto sistematico di milioni di persone ai campi di concentramento, nonché il loro sterminio. Ma la Seconda Guerra Mondiale è stata combattuta, e milioni di persone sono stati sterminati, quando il computer ancora non c'era.

nel fatto che essa ci sensibilizza sulla possibilità che certe azioni umane, per esempio l'introduzione di computer in attività umane complesse, costituiscano una presa d'impegno irreversibile. Non è vero che il sistema bancario o quello commerciale americano, o le grandi aziende produttrici, sarebbero crollati se il computer non fosse arrivato «appena in tempo». È vero invece che il modo particolare in cui questi sistemi si sono poi sviluppati durante gli ultimi due decenni, e in cui si stanno sviluppando tuttora, non avrebbe potuto essere in assenza del computer. È vero che, se tutti i computer scomparissero improvvisamente, gran parte del mondo industrializzato e militarizzato moderno si ritroverebbe immerso in una grande confusione, se non nel caos più completo. Il computer non era un pre-requisito per la sopravvivenza della società moderna nel periodo post-bellico e in quello successivo; l'accoglienza entusiastica e acritica riservatagli dagli esponenti più «progressisti» del governo, del mondo degli affari, dell'industria americani ne ha rapidamente fatto una risorsa essenziale per la sopravvivenza della società *in quella forma* che il computer stesso, come strumento, aveva contribuito a determinare.

Nel 1947 J.W. Forrester scrisse un memorandum indirizzato alla Marina degli Stati Uniti, *Sull'uso dei calcolatori digitali elettronici come centri d'informazione per il combattimento.* Commentando gli eventi successivi, nel 1961 egli ha scritto:

> [...] probabilmente (nel 1947) sarebbe stato impossibile trovare cinque ufficiali disposti ad accettare la possibilità che una macchina fosse in grado di analizzare le fonti d'informazione disponibili, la migliore assegnazione delle armi, l'emissione di istruzioni e comandi, e il coordinamento di aree vicine di operazioni militari [...] Durante il decennio successivo, la rapidità delle operazioni militari aumentò finché divenne chiaro che, indipendentemente dai presunti vantaggi dell'affidare all'uomo le decisioni, la velocità di comunicazione dell'organizzazione umana semplicemente non era in grado di sostenere il ritmo delle operazioni militari aeree moderne. Questa incapacità costituì l'incentivo [9].

Il decennio cui Forrester si riferisce fu pieno di incentivi simili, di scoperte del fatto che l'organizzazione umana esistente stava avvicinandosi a dei limiti nella sua capacità di affrontare il ritmo sempre più veloce della vita moderna. L'immagine evocata da Forrester è quella di piccoli gruppi di uomini che si affannano cercando di tener dietro agli eventi, ma che rimangono sempre più indietro perché gli avvenimenti si susseguono troppo rapidamente, e c'è troppo da fare. Quegli uomini hanno raggiunto il limite della «massima velocità interna» del gruppo. La stessa immagine può forse essere usata per descrivere provocatoriamente gruppi di impiegati di banca che registrano e controllano assegni nel mezzo della notte, con davanti montagne sempre più alte di assegni che, secondo la legge, devono essere registrati entro una scadenza

[9] J.W. Forrester, in M. Greenberger (a cura di), *Managerial Decision Making in Management and the Computer of the Future* («Il processo decisionale manageriale e il computer del futuro»), The MIT Press, Cambridge, Mass. 1962, pp. 52-53.

fissata. Molti dei limiti, di altro tipo, che vennero raggiunti durante quel decennio potrebbero essere descritti in modo simile. Dopo tutto, è in fin dei conti la «velocità interna» di qualche organizzazione umana a costituire il fattore limitante quando, poniamo il caso, una fabbrica di automobili tenta di impostare una linea di produzione in grado di produrre una varietà astronomica di automobili a un ritmo elevato e costante o quando, poniamo, una qualche istituzione governativa si assume la responsabilità di sorvegliare milioni di cittadini che usufruiscono di assistenza — affinché non tentino di truffare —, controllando dettagliatamente le sovvenzioni che ricevono e ogni altra loro forma di entrata, magari illecita.

La «incapacità di agire» che, come osserva Forrester, «ha fornito l'incentivo» a integrare o sostituire l'organizzazione umana (a bassa velocità interna) con il computer, avrebbe potuto essere, in qualche altra situazione storica, un incentivo per modificare l'obiettivo da raggiungere, magari rinunciandovi del tutto, o per ristrutturare l'organizzazione umana le cui limitazioni intrinseche, dopo tutto, erano alla radice del problema. Forse, l'incentivo fornito dall'incapacità dei militari di sostenere la crescente complessità delle operazioni aeree negli anni Cinquanta avrebbe potuto essere tradotto in una preoccupazione: non quella di trovare tecniche che mettessero i militari in grado di adempiere ai loro incarichi tradizionali, bensì quella di inventare nuove organizzazioni umane con nuovi incarichi, incarichi relativi a questioni più fondamentali, per esempio come far convivere popoli con interessi diversi. Ma il computer venne usato per costruire, nelle parole di un colonnello d'aviazione, «un servomeccanismo sparso su una superficie paragonabile a quella di tutto il continente americano», vale a dire il sistema di difesa aerea SAGE. Naturalmente, una volta entrati in possesso di un simile sistema, dovemmo supporre che anche «loro» ne avessero uno. Dovemmo perciò applicare la nostra tecnologia al progetto di armi e strategie in grado di superare la «nostra» difesa, cioè la presunta difesa «loro». Quindi dovemmo supporre che «loro» avessero armi e strategie simili, e perciò... fino agli odierni MIRV e MARV e ABM.

Forse, la fame di automobili private — che è stata ben coltivata e ha finito per dare assuefazione — avrebbe potuto essere saziata dando alla gente la possibilità, mettiamo, di scegliere tra cento veicoli davvero sostanzialmente diversi uno dall'altro, piuttosto che tra un numero enorme di «modelli» sostanzialmente identici, che differiscono solo per dettagli banali. Anzi, forse l'automobile privata avrebbe potuto essere penalizzata, in quanto mezzo di trasporto personale, a favore del trasporto pubblico nelle città, e dei collegamenti ferroviari tra le città. Ma il computer è stato invece usato per automatizzare il flusso di componenti nelle catene di montaggio, così che oggi la gente può scegliere tra milioni di insignificanti *optional* sulla macchina nuova.

Forse, servizi sociali quali l'assistenza pubblica avrebbero potuto essere amministrati da esseri umani, con la loro umana capacità di giudicare, se tali servizi fossero stati organizzati in modo decentrato, sulla scala delle comunità locali. Ma il computer è stato invece usato per automatizzare l'amministrazione dei servizi sociali, e per centralizzarli secondo linee politiche prestabilite. Se il computer non avesse facilitato il perpetuarsi e l'«ammodernamento» dei

sistemi di distribuzione dell'assistenza già esistenti — e quindi del loro fondamento filosofico —, magari qualcuno avrebbe potuto pensare di eliminare gran parte del bisogno di assistenza, per esempio introducendo una tassa negativa sul reddito. Il fatto stesso di aver creato un apparato per l'amministrazione dell'assistenza enormemente esteso e complicato, e basato sul computer, ha creato un interesse al suo mantenimento, e quindi la perpetuazione del sistema stesso. Questo interesse è presto diventato una barriera di fronte a ogni innovazione, anche se in seguito le buone ragioni per innovare si accumularono. In altre parole, molti dei problemi di crescita e di complessità che richiesero insistentemente e irresistibilmente una risposta, nei decenni post-bellici, avrebbero potuto fungere da incentivi per un rinnovamento sociale e politico. Se in quegli anni fosse iniziata una grande accelerazione dell'inventiva legata ai problemi sociali, essa ci sembrerebbe oggi una conseguenza naturale dell'atteggiamento dell'uomo in quel periodo, così come naturale ci sembra il grande flusso di invenzioni e rinnovamenti verificatisi in campo tecnologico.

Sì, il computer è arrivato «appena in tempo». Ma in tempo per cosa? In tempo per salvare — e salvare quasi intatte, anzi riassestare e stabilizzare — strutture sociali e politiche che avrebbero altrimenti potuto essere o rinnovate radicalmente, o lasciate vacillare di fronte allo sforzo che certamente sarebbe stato loro richiesto. Il computer, quindi, è stato usato per preservare le istituzioni sociali e politiche americane. Le ha vaccinate e immunizzate, almeno temporaneamente, contro enormi spinte verso il cambiamento. La sua influenza è stata sostanzialmente la stessa in altre società che hanno permesso al computer di fare strada all'interno delle loro istituzioni: vengono immediatamente in mente il Giappone e la Germania.

L'invenzione del computer ha messo in pericolo una parte di un mondo apparentemente stabile, così come è compito di ogni azione creativa dell'uomo. E, proprio come diceva Dewey, nessuno avrebbe potuto prevedere cosa sarebbe emerso al posto di quella parte. Ma delle molte strade verso l'innovazione sociale che il computer ha aperto all'uomo, la più decisiva è stata metterlo in grado di neutralizzare ogni deliberata idea di cambiamento radicale. Ed è questa la strada che l'uomo ha scelto di seguire. L'arrivo della *Rivoluzione del Computer* e la fondazione dell'*Era del Computer* sono stati annunciati molte volte. Ma se il trionfo di una rivoluzione va misurato in termini di profondità dei cambiamenti sociali che ha determinato, allora non c'è stata alcuna rivoluzione del computer. E qualunque sia il modo giusto di caratterizzare l'era attuale, è certo che il computer non merita di darle il nome.

Dire che il computer è stato usato inizialmente per fare certe cose nel modo in cui erano sempre state fatte, però più rapidamente o — secondo un qualche criterio — in modo più efficiente, non significa distinguerlo dagli altri strumenti. Solo raramente, ammesso che succeda mai, uno strumento è inventato insieme a un lavoro del tutto originale da fare. Gli strumenti come simboli, tuttavia, tendono naturalmente a farsi spostare in contesti diversi dai loro originali. Nel loro nuovo sistema di riferimento, vale a dire come nuovi simboli in un disegno immaginario già stabilito, possono essere trasformati essi stessi, e possono addirittura trasformare le prescrizioni del disegno originario. Que-

ste trasformazioni possono, a loro volta, creare problemi completamente nuovi, che finiscono per generare l'invenzione di strumenti fino ad allora letteralmente inimmaginabili. Nel 1804, cent'anni dopo che il primo motore a vapore stazionario di Newcomen e Savery era divenuto di uso comune in Inghilterra, per esempio per pompare acqua fuori dalle miniere, Trevithik mise un motore a vapore su un carro, e il carro sui binari di una tramvia a cavalli in Galles. Questo sradicamento del motore a vapore stazionario dal suo contesto, e il suo spostamento in un contesto completamente nuovo, trasformarono il motore in una locomotiva, e diedero inizio alla trasformazione delle tramvie a cavalli nella moderna ferrovia. E tra le altre cose, siccome diventò necessario evitare le collisioni fra treni viaggianti sullo stesso binario, venne stimolato lo sviluppo di una tecnologia della segnalazione completamente nuova. Erano stati creati nuovi problemi e, come risposta, inventati nuovi strumenti.

È degno di nota il fatto che Thomas Savery, il costruttore del primo motore a vapore applicato alla produzione industriale (1700 circa), fu anche il primo ad usare l'espressione «cavallo-vapore» («HP», *horsepower*) più o meno nella sua accezione moderna. Forse il termine nacque soltanto perché il cavallo era così comune quando il motore a vapore cominciò a sostituirlo, non solo nella sua prima incarnazione di fonte stazionaria di energia, ma anche nella sua reincarnazione di locomotiva. Eppure l'espressione «cavallo-vapore», così ricca di suggestione, potrebbe anche aver spinto la fantasia di Trevithik nella direzione che egli infine prese, verso il salto creativo che combinò motore a vapore e tramvia a cavalli in un unico sistema di riferimento. L'invenzione comporta una proiezione fantasiosa di simboli da un sistema di riferimento esistente, e in genere ben sviluppato, a un altro. C'è da aspettarsi che alcuni simboli particolarmente potenti sopravviveranno quasi intatti al passaggio, e finiranno per esercitare la loro influenza anche sul nuovo sistema di riferimento.

I computer dovettero sostituire un altro tipo di cavalli. Prima dell'entrata in uso dei calcolatori digitali elettronici moderni, per quella che oggi chiamiamo elaborazione dati commerciali — vale a dire, prima dell'acquisto di UNIVAC I da parte dell'ufficio del censo degli Stati Uniti nel 1951 —, in molte imprese americane funzionavano grandi cosiddette *tab rooms* (stanze per la produzione di tabelle). Queste stanze ospitavano macchine capaci di interpretare delle schede, del tipo di quelle ancora in uso fino a pochi anni fa (chiamate comunemente, anche se erroneamente, «schede IBM»), in base a criteri arbitrari, e di «tabulare» infornate di tali schede, elencando cioè in lunghe tabelle stampate l'informazione in esse contenuta. Le *tab rooms* produssero montagne di rapporti di gestione per il governo e l'industria americani, usando metri cubi di enormi mostri meccanici sferraglianti. Queste macchine potevano eseguire soltanto un'operazione per volta su ogni blocco di schede. Per esempio, potevano analizzare il blocco sulla base di una specifica chiave di analisi. Se il blocco così analizzato richiedeva un ulteriore scrutinio sulla base di un altro criterio, quest'ultimo andava inserito manualmente nella macchina e il blocco di schede doveva essere fatto analizzare un'altra volta. Le *tab rooms* erano le tramvie a cavalli dell'elaborazione di dati commerciali, le *tab machines* erano i cavalli.

In linea di principio anche il primissimo calcolatore elettronico messo in commercio, l'UNIVAC I, rese possibili tecniche di elaborazione di dati completamente nuove e molto più efficaci proprio come, in linea di principio, il primissimo motore a vapore avrebbe potuto essere montato su un carro, e il carro su una rotaia. Per la verità, durante e subito dopo la Seconda Guerra Mondiale l'arte della ricerca operativa e dell'analisi dei sistemi, sulla quale venne basato in seguito l'uso sofisticato del computer a fini commerciali, venne sviluppata fin quasi alla completa maturità. E tuttavia, il mondo degli affari usò i primi computer semplicemente per «automatizzare» le sue *tab rooms,* cioè per svolgere esattamente le stesse operazioni di prima, soltanto in modo automatico e, presumibilmente, più efficiente. La transizione cruciale, dal computer semplice sostituto delle primordiali macchine per tabelle alla sua attuale condizione di versatile macchina informatica, iniziò quando la potenza del computer venne proiettata sull'ordito già elaborato dalla ricerca operativa e dall'analisi dei sistemi.

È bene aggiungere qui che, anche se la ferrovia in Inghilterra divenne importante di per sé — diede lavoro a molte persone, per esempio —, essa accrebbe anche enormemente l'importanza di molti altri mezzi di trasporto. In modo simile, la combinazione sinergetica di computer e analisi dei sistemi ebbe un ruolo fondamentale nella creazione e nella crescita dell'industria elettronica. E stimolò anche una nuova vitalità nell'analisi dei sistemi in quanto tale. Durante il primo decennio della sistematica invasione del mondo degli affari da parte del computer, quando accadeva spesso che un manager decidesse di avere bisogno del computer avendo soltanto una vaga idea di cosa ciò significasse, succedeva anche che egli intraprendesse un'approfondita analisi di sistemi sulle sue necessità, per stabilire che cosa dovesse fare il nuovo computer. In moltissimi casi, quest'analisi rivelava la possibilità di migliorare le operazioni, talvolta radicalmente, senza introdurre affatto il computer. Né il computer veniva usato per lo studio stesso. Spesso, naturalmente, il computer veniva installato lo stesso per ragioni, diciamo, di moda o di prestigio.

Un effetto collaterale di questa reiterata esperienza fu il deciso riconoscimento dell'analisi dei sistemi, e in misura minore dell'analisi delle operazioni, come metodologia per prendere decisioni in campo commerciale. Siccome il prestigio dell'analisi dei sistemi era rafforzato dai suoi successi e, simultaneamente, il computer cresceva d'importanza, i problemi affrontati dagli analisti di sistemi divennero sempre più complessi, e il computer apparve sempre più come lo strumento idoneo a trattare la complessità. Di solito l'analisi dei sistemi, almeno all'osservatore casuale, dà l'impressione di essere stata fagocitata dal computer. Questa apparenza è fuorviante, ma non priva di significato. L'analisi dei sistemi è sopravvissuta e ha prosperato come disciplina a sé stante. Il computer ha dato muscoli alle sue tecniche. Le ha talmente rafforzate da renderle qualitativamente diverse dalle tecniche manuali originarie. Queste ultime, di conseguenza, sono in gran parte scomparse. E quelle nuove non possono più venir considerate separatamente dal computer.

L'interazione del computer con l'analisi dei sistemi è istruttiva anche da un altro punto di vista. È importante capire molto chiaramente che il rafforza-

re una particolare tecnica — il darle muscoli potenti — non contribuisce affatto alla sua validità. Per esempio, ci sono programmi per computer che eseguono con grande precisione tutti i calcoli necessari a fare l'oroscopo di una persona di cui si conosca l'ora e il luogo della nascita. Poiché è il computer a fare tutte le noiose manipolazioni di simboli, l'oroscopo può venire prodotto molto più rapidamente di quanto sia possibile normalmente a un astrologo. Ma questo miglioramento nella tecnica degli oroscopi è del tutto irrilevante rispetto alla questione della validità delle previsioni astrologiche. Se l'astrologia è una sciocchezza, allora quella computerizzata lo è altrettanto. Ora, talvolta certe semplici tecniche sono prive di validità nel loro campo d'applicazione soltanto a causa della loro semplicità, mentre tecniche dello *stesso* tipo, ma molto più complicate, sono valide in quel campo. Ciò non vale per l'astrologia, ma può essere vero, ad esempio, per le previsioni del tempo fatte usando il calcolatore. In questo caso, il numero dei dati da prendere in considerazione e la massa di calcoli da farci sopra per ottenere un'accurata previsione del tempo possono benissimo essere così grandi che nessun gruppo di uomini, per quanto numeroso, potrebbe avere risultati in un tempo ragionevole. E qualunque semplificazione della tecnica, tale da ridurre i problemi di calcolo a dimensioni affrontabili dall'uomo, invaliderebbe la tecnica stessa. In casi come questo, il computer può contribuire a rendere fattibile una tecnica fino ad allora impraticabile. Ma ciò che non va dimenticato è che la validità di una tecnica è un problema che coinvolge la tecnica stessa e anche il campo in cui la si applica. Se si vuole trasformare una cattiva idea in una buona, è la *fonte* della sua debolezza che bisogna scoprire e correggere. È improbabile che si riesca ad aiutare un uomo che cade in un tombino dandogli la possibilità di cadere in modo più veloce o efficiente.

Può sembrare strano, o anche paradossale, che il potenziamento di una tecnica possa metterne in mostra debolezze e limitazioni, ma questo non dovrebbe sorprenderci. La mente umana ha una grande capacità di pensiero contorto e di razionalizzazione, vale a dire, giustificazione in termini apparentemente razionali delle conseguenze dei suoi pensieri contorti. Se una particolare tecnica richiede una gran quantità di calcoli, mentre ad essa si può dedicare solo una modesta quantità di tempo, allora un insuccesso della tecnica può facilmente essere spiegato sostenendo che essa, a causa dei limiti computazionali, non è stata mai provata sul serio. La tecnica stessa è immunizzata contro un esame critico, grazie a queste scappatoie. Anzi, può finire per esserne fortificata, perché il credere che una tecnica altrimenti perfetta e probabilmente molto potente sia invalidata da qualche limitazione isolata, tende a spingere il fanatico a fare ogni sforzo per eliminare quella limitazione. Se essa gli sembra essere interamente computazionale, e se è disponibile un computer, egli potrà finire per lanciare una campagna di «ricerca intensiva», che richiede molto tempo-macchina, col solo fine di «computerizzare» la sua tecnica. Programmi di questo tipo generano di solito sotto-problemi di natura strettamente computazionale che tendono, proprio in virtù delle loro dimensioni, a dominare sempre più l'impresa e, se non si sta più che attenti, a diventare infine il centro dell'attenzione. E mentre si investe sempre di più per affrontare questi proble-

mi inizialmente secondari, e si progredisce nel risolverli, tende a crescere l'illusione di stare lavorando in modo serio sul problema principale. La povertà della tecnica, nel caso in cui sia davvero impotente di fronte ai problemi che dovrebbe risolvere, viene così nascosta dietro una montagna di lavoro, molto del quale può anche avere successo nel suo settore particolare. Ma tale settore fa parte di un contesto determinato dal problema cui la tecnica originaria era stata applicata. L'insieme dei sotto-problemi con le convenzioni, il gergo, le sottotecniche cristallizzate intorno a loro, diventa così realtà. Quanto più grande è questo insieme, quanta più energia umana è stata investita nella sua creazione, tanto più esso sembra reale. E quanto più difficile è stato risolvere i sottoproblemi, quanto più esaltante è stato risolverli dal punto di vista tecnico, tanto più la tecnica originale si ritrova ad essere fortificata.

Ho discusso il ruolo degli strumenti nella ricostruzione fantasiosa che l'uomo fa del suo mondo, e nell'affinamento delle sue tecniche. Tuttavia, gli strumenti hanno anche un altro ruolo: costituiscono una sorta di linguaggio per la società che li adopera, un linguaggio di azione sociale. Più avanti parlerò ancora del linguaggio. Per ora, accontentiamoci di caratterizzare il linguaggio, in modo un po' incompleto, come consistente in un vocabolario — le parole del linguaggio — e in un insieme di regole che determinano come i singoli vocaboli possano essere concatenati a formare frasi dotate di significato. Per il momento, lascio da parte gli innumerevoli misteri che circondano il concetto di «significato». Mi limito all'accezione più stretta, vale a dire quello di azione che una particolare «frase» nel linguaggio degli strumenti inizia e svolge.

Il linguaggio ordinario deriva il suo potere espressivo in parte dal fatto che ogni sua parola ha un campo ristretto di significato. Sarebbe impossibile dire qualunque cosa in un linguaggio che consistesse interamente di pronomi, per esempio. Anche strumenti e utensili derivano il loro potere dal fatto di permettere certe azioni e non altre. Per esempio, un martello deve essere rigido. Non può dunque essere usato come fune. Non può esistere un utensile tuttofare, così come non può esistere una parola che significa tutto. Sappiamo bene che l'uso di parole specifiche in un'accezione generalizzata, come accade per «cioè» e «appunto», ha l'effetto di impoverire — anziché arricchire — la lingua.

Forse, inventare strumenti nuovi è tanto difficile quanto inventare parole veramente nuove. Tuttavia, il XX secolo ha assistito all'invenzione almeno di un modesto numero di strumenti che estendono davvero il possibile raggio d'azione della società. Automobili e autostrade, radio e televisione, medicine e tecniche chirurgiche moderne sono i primi che vengono alla mente. Tutte queste cose hanno messo la società in grado di articolare attività mai possibili prima. Quello che viene detto di rado, tuttavia, è che i nuovi modi di agire della società spesso eliminano la possibilità stessa di agire nei modi vecchi. Qualcosa di analogo succede nel linguaggio ordinario. Per esempio, ora che l'espressione «non operativo» è stata usata da alti ufficiali governativi americani come eufemismo per «mentire», essa non può più essere usata per comunicare il significato originario. Termini come «libero» (come in «il mondo libero»), «soluzione finale», «difesa» e «aggressione» sono stati così profondamente degradati da un uso corrotto, da diventare essenzialmente inutili per la conversa-

zione ordinaria. In modo simile, un'autostrada permette alla gente di viaggiare tra i centri geografici che collega ma, a causa degli effetti collaterali che essa genera sinergeticamente insieme ad altri fattori, finisce per imprigionare la povera gente delle città minori come se tali città fossero murate. I mezzi di comunicazione di massa, si dice talvolta, hanno ridotto la Terra a un «villaggio globale», e hanno reso possibili incontri nazionali, o addirittura «assemblee civiche» globali. Ma, in contrasto con le assemblee civiche tradizionali del New England, che erano — e rimangono, nella mia città — un esercizio di politica della *partecipazione*, i *mass media* sostanzialmente non permettono al cittadino di rispondere. Come le autostrade e le automobili, essi consentono alla società di articolare forme completamente nuove di azione sociale, ma allo stesso tempo disattivano irreversibilmente certe modalità di comportamento sociale che erano disponibili prima.

Il computer, in un certo senso, è uno strumento di questo tipo. Ha aiutato l'uomo a socchiudere la porta verso lo spazio esterno, e ha salvato certe istituzioni sociali minacciate di crollo sotto il peso di una popolazione in rapida crescita. Ma il suo impatto ha anche chiuso certe porte che una volta erano aperte... se questo sia un processo irreversibile o meno, non possiamo ancora dirlo con certezza. C'è un mito secondo cui i computer oggi prendono decisioni importanti, che in precedenza erano prese dall'uomo. Ci sono forse esempi isolati di questo, qui e là nella nostra vita sociale. Ma l'immagine, molto diffusa, del manager che batte sulla tastiera: «E ora cosa facciamo?» e aspetta che il computer «decida», è in gran parte sbagliata. Quello che sta succedendo, invece, è che la gente ha affidato a sistemi di computer enormemente complicati l'elaborazione delle informazioni su cui le decisioni devono basarsi. Ma la gente, con poche eccezioni, tiene per sé il diritto di prendere le decisioni sulla base del risultato di quei calcoli. In questo modo, gli uomini possono conservare l'illusione, e spesso è proprio solo un'illusione, di essere dopotutto loro a prendere le decisioni. Ma, come discuteremo, un sistema di calcolo che permette di fare solo certi tipi di domande, che accetta solo certi tipi di «dati», e che non può essere compreso — nemmeno più in linea di principio — da coloro che gli si affidano, un sistema di questo tipo ha in effetti chiuso molte porte che erano aperte prima della sua installazione.

Per capire in che modo il computer abbia raggiunto un simile potere, sia come attore autonomo che come forza che preme sulla fantasia umana, dobbiamo innanzitutto discutere la provenienza di tale potere, e come il computer riesca a fare quello che fa. È a questi due punti che dedicheremo la nostra attenzione nei prossimi due capitoli.

2
DA DOVE VIENE IL POTERE DEL COMPUTER

Se ci capitasse di vedere qualcosa che per noi è molto strano, diciamo una nuvola con un contorno molto netto e ben delineato, vorremmo sapere che cos'è. E se ci dicessero che è una fuba, chiederemmo che cos'è una fuba. Ma intorno a noi ci sono degli oggetti che fanno così costantemente parte della nostra vita che non li troviamo strani, né chiediamo che cosa siano. Succede così con le macchine. La parola «macchina» ci richiama alla mente immagini di movimenti complessi, eppure in qualche modo regolari. Il movimento dell'ago di una macchina per cucire, così simile a quello delle barre di connessione che spingono le ruote di una locomotiva, o al tremito del meccanismo pulsante del più delicato orologio da polso: sono immagini che quasi assommano in sé la nostra idea di «macchina». Quasi. Quanto basta per non farci più domandare che cosa sia una macchina. Regolarità, complessità, movimento, potenza. Eppure c'è qualcosa di più, e lo sappiamo.

Mettiamo in moto una pressa, e questa stritola la mano di chiunque le si avvicini troppo. La caratteristica più paurosa delle macchine è proprio la loro regolarità. Quando sono in moto fanno il loro lavoro, certamente in modo regolare, ma anche in modo cieco. Quando diciamo che la giustizia è cieca, intendiamo dire che essa deve essere simile ad una macchina, che svolge la sua attività senza curarsi di fatti irrilevanti — pur sempre fatti, comunque. Per la giustizia cieca, il fatto che il prigioniero alla sbarra sia povero o ricco, uomo o donna, è irrilevante. Per la pressa, il fatto che tra le sue ganasce ci sia un pezzo di metallo o la mano di un operaio è irrilevante. Come tutte le macchine, la giustizia cieca e la pressa fanno ciò che sono destinate a fare; e lo fanno in modo esatto.

Le macchine, quando lavorano a dovere, non obbediscono solamente a certe leggi; sono *incarnazioni* di quelle leggi. Dire che una macchina «funziona a dovere» equivale a dire che essa è un'incarnazione di una legge che conosciamo e che vogliamo applicare. Per esempio, ci aspettiamo che una normale

[1] I capitoli 2 e 3 sono piuttosto tecnici. Il lettore che non ama gli aspetti tecnici può saltarli, oppure leggerli dopo aver letto il resto del libro.

calcolatrice da tavolo sia un'incarnazione delle leggi dell'aritmetica che tutti conosciamo. Se ci dà un risultato che crediamo sbagliato, la nostra fede nella obbedienza della macchina alle leggi è così forte che, di solito, immaginiamo di aver fatto noi un errore nell'introdurre i dati. Soltanto quando gli errori si ripetono, decidiamo che «qualcosa non va nella macchina». Non crederemmo mai che le leggi dell'aritmetica abbiano cessato di valere o siano state modificate. Ma non siamo neppure disposti a credere che la macchina si comporti in modo capriccioso, rifiutandosi cioè di obbedire alle leggi. No, per farla funzionare di nuovo bene cerchiamo di capire perché si comporti come sta facendo *ora*, vale a dire, di quale legge sia ora un'incarnazione. Ci fa piacere trovare, mettiamo, una rotella rotta che *giustifichi* il suo comportamento aberrante. In tal caso, abbiamo scoperto quella legge. Siamo ora in grado di capire la macchina che abbiamo tra le mani, e di ripararla, vale a dire di trasformarla nella macchina che avevamo all'inizio, nell'incarnazione delle consuete leggi dell'aritmetica. E di solito ci dà molto fastidio se un tecnico ci restituisce una macchina dicendo «Non so che cosa non andasse: l'ho solo aperta e chiusa, e adesso funziona alla perfezione». Egli confessa di non essere riuscito a capire la legge della macchina rotta, e noi ne deduciamo che egli ora non può sapere, né lo possiamo noi, quale sia la legge della macchina «riparata». Se noi dipendiamo da quella macchina, siamo diventati servi di una legge che non possiamo conoscere, quindi di una legge arbitraria, capricciosa. È questa la ragione del nostro disappunto.

Le macchine che popolano il nostro mondo non sono più esclusivamente, anzi neppure principalmente, dei mostri sferraglianti che vengono riconosciuti come macchine proprio in virtù del rumoroso movimento delle loro parti. Esistono oggi degli orologi i cui ingranaggi sono circuiti stampati su piccolissimi rettangoli di plastica, orologi del tutto privi di parti in movimento. Anche le lancette sono scomparse. Ci dicono l'ora, quando gliela chiediamo, facendo apparire dei numeri luminosi sul loro quadrante. Le bobine rotanti che in passato distribuivano scariche elettriche alle candele dei motori delle nostre automobili sono state rimpiazzate da piccole scatole nere, che di nuovo contengono circuiti stampati su minuscoli *chip* di plastica, che silenziosamente e senza movimento alcuno producono gli impulsi richiesti. Anche questi, come migliaia di altri apparecchi simili, li chiamiamo macchine.

Questa estensione del dominio connotativo della parola «macchina» ha due significati piuttosto indipendenti: in primo luogo, testimonia che il buon senso comune riconosce come caratteristica essenziale di una macchina la sua infaticabile regolarità, la sua cieca obbedienza alla legge che incarna. E tale regolarità, come il buon senso comune correttamente riconosce, ha poco a che vedere con l'effettivo movimento. È questa l'intuizione che permette alla gente di parlare, mettiamo, di una burocrazia o di un sistema giudiziario come di una macchina. In secondo luogo, essa rivela che il buon senso comune afferra implicitamente, ancorché in modo molto vago, l'idea che il meccanismo ha in parte a che fare con il trasferimento di informazione, piuttosto che con la trasmissione di potenza in senso stretto. La comparsa di macchine elettroniche di tutti i tipi, e in particolare del calcolatore elettronico, ha modificato la

nostra immagine della macchina da quella di un conduttore e trasmettitore di *potenza* a quella di un trasformatore di *informazione*.

Molte altre macchine hanno componenti interni la cui funzione primaria è trasmettere informazione benché, nel complesso, la funzione di tali macchine sia fornire potenza meccanica. Considerate, per esempio, un comune motore a benzina a quattro tempi. Esso è, naturalmente, un generatore di potenza. Uno dei suoi componenti è la biella, una sbarra di acciaio la cui estremità inferiore corre su un albero a camme, mentre quella superiore può aprire la valvola di scarico del cilindro cui appartiene. Quando il motore fa girare l'albero principale, fa girare anche l'albero a camme e la camma su cui corre la biella. Questa, perciò, esegue un movimento in su e in giù che apre e chiude la valvola di scarico del cilindro proprio al momento giusto. Nei più semplici motori a benzina, la biella fornisce sia la potenza per far muovere la valvola che il sincronismo necessario. In motori più complicati, però, essa funge semplicemente da segnalatore per qualche altro meccanismo che manipola la valvola. Possiamo quindi immaginare di sostituirla con un circuito connesso, a un estremo, con un apparecchio che avverte quando è necessario espellere gas da un cilindro e, all'altro estremo, con un motore che apre e chiude opportunamente la valvola. Molti moderni motori da automobile sono dotati di sistemi elettronici di iniezione del carburante che, grosso modo, funzionano proprio così.

C'è, tuttavia, un limite al numero di connessioni meccaniche in un motore di automobile che possono essere sostituite da apparecchi che trasmettono informazione. Il motore, dopo tutto, deve distribuire potenza meccanica alle ruote. Questo requisito pone severi limiti al progettista del motore. Un ingegnere può benissimo concepire un insieme di leggi coerente, in altre parole un progetto, per un motore che ciononostante non può essere realizzato. Il suo progetto potrebbe richiedere la lavorazione di metalli fino a livelli di tolleranza non raggiungibili con le tecniche che egli ha a disposizione, per esempio. Oppure, la robustezza dei materiali richiesti dal suo motore può essere irrealizzabile con le tecnologie dell'epoca. Ma, cosa molto più importante, il suo progetto può essere irrealizzabile in linea di principio, perché viola qualche legge fisica. È questo, per esempio, lo scoglio con cui si scontreranno tutte le macchine basate sul moto perpetuo. Le leggi incarnate da una macchina che interagisce con il mondo reale devono, per forza di cose, essere un sottoinsieme delle leggi che governano il mondo reale.

È, naturalmente, insensato parlare di una macchina, fatta di sostanze materiali, che non interagisca con il mondo reale. Se una cosa simile esistesse, non potremmo conoscerla (in quanto, per farci sapere della sua esistenza, dovrebbe interagire con i nostri sensi, cioè con il mondo reale). In ogni caso, una macchina di questo tipo sarebbe inutile: per noi «utilità» non significa forse interazione con il mondo?

Ma ci sono situazioni in cui è lecito parlare di aspetti di macchine reali che sono separati dall'incarnazione fisica della macchina. Talvolta, per esempio, ci serve discutere di cosa una macchina — o una sua parte — debba fare, senza preoccuparci di come, o con quali materiali, sarebbe possibile costruire

davvero un apparecchio che svolga la funzione desiderata. Per esempio, nel motore a benzina ci deve essere una parte che riconosce quando la valvola di scarico di un cilindro deve essere aperta, e quando chiusa. Tale funzione può essere svolta da una biella o, come ho detto, da un filo che collega opportunamente un sensore e un motorino elettrico. La legge che tale strumento deve seguire, quella di cui è un'incarnazione, è un'idea astratta. È indipendente dalla materia, dall'incarnazione materiale, in poche parole da tutto ciò che non è pensiero e ragionamento. Da una regola simile — o «specificazione funzionale», come amano dire gli ingegneri — si possono ricavare molti progetti diversi; per esempio, uno può consistere in una biella meccanica, un altro in una «biella elettrica». Il progetto di una macchina è, di nuovo, un'astrazione. Un buon progetto, mettiamo, di una macchina per cucire potrebbe venire usato da diverse fabbriche, ciascuna delle quali produrrebbe macchine per cucire sostanzialmente indistinguibili da quelle di un'altra. In un certo senso, quindi, quel progetto è una macchina per cucire astratta. È la macchina per cucire che potrebbe venire prodotta, meno, per così dire, i componenti materiali — l'*hardware* — della macchina per cucire vera e propria. Il progetto è anche indipendente dal modo in cui viene messo nero su bianco. Il disegno di una macchina non è il suo progetto. Se lo fosse, allora il progetto cambierebbe ogni volta che il disegno viene ingrandito, o ridisegnato in un altro colore. No, un progetto è un'idea astratta, così come lo è una specificazione funzionale. E le idee — per esempio l'idea di una macchina a moto perpetuo — non sono vincolate dalle leggi della fisica.

Gli scrittori di fantascienza saltano sempre fuori, in effetti, con specificazioni funzionali per macchine che sono di solito fisicamente irrealizzabili, in quanto violerebbero principi fisici insormontabili. Un'idea che emerge ogni minuto è quella della comunicazione istantanea tra luoghi molto distanti. La fisica, però, sostiene che nessun messaggio, di qualunque tipo, può venire inviato da un luogo ad un altro a una velocità superiore a quella della luce. E dal momento che la velocità della luce è finita (circa 300.000 chilometri al secondo), la comunicazione istantanea, anche su distanze brevi, è impossibile — almeno secondo la fisica moderna. Allora, le idee presentate dagli scrittori di fantascienza sono inutili? No. Perché, anche se il nostro corpo deve funzionare in un mondo vincolato dalle leggi della natura, la nostra mente ne può scappare. Possiamo dare libero sfogo alle nostre idee in un mondo fittizio, in cui non esista barriera alla velocità di comunicazione. Questo equivale a dire che possiamo giocare a giochi di cui stabiliamo noi stessi le regole. E possiamo scegliere fino a che punto le regole del nostro gioco debbano corrispondere alle leggi che — crediamo — governano il mondo reale. Il gioco «Monopoli» potrebbe esistere anche in mondi in cui l'avidità non è un dato di fatto (se ce ne fossero).

Una proprietà fondamentale che l'insieme delle regole di qualsiasi gioco deve avere è quella di essere completo e coerente. Deve essere completo nel senso che, di fronte a qualsiasi proposta di azione all'interno del gioco, le regole devono permettere di decidere se quell'azione è lecita oppure no. Deve essere coerente nel senso che non deve accadere che un sottoinsieme delle rego-

le dica che una certa azione è lecita, mentre allo stesso tempo un altro sottoinsieme sostiene che la stessa azione non è lecita [2].

Un gioco puramente astratto è un gioco le cui regole non implicano alcun contatto con il mondo reale, cioè un gioco che può essere giocato col solo ragionamento. Le partite di scacchi dei tornei non sono un gioco puramente astratto, perché il regolamento prevede un limite al tempo che un giocatore ha a disposizione per decidere le sue mosse. Questo riferimento al tempo mette gli scacchi in contatto con il mondo reale, rovinando così la purezza della loro astrazione. Se si dimentica questo vincolo, comunque, gli scacchi sono un gioco puramente astratto. C'è un altro modo per esprimere la condizione che le regole di un gioco devono essere complete e coerenti, nel senso che si intende qui: non deve essere possibile che due arbitri, trovandosi di fronte alla stessa situazione di gioco, esprimano giudizi differenti. Anzi, «giudizio» non è neppure la parola corretta da usare, perché la loro decisione verrebbe presa usando soltanto la logica. Sarebbe, in realtà, niente più di un preciso calcolo, un processo logico che può avere un solo risultato.

C'è un solo tipo di problema che potrebbe essere soggetto al giudizio di un arbitro, in un gioco puramente astratto. Un giocatore potrebbe descrivere la situazione di gioco, per esempio la posizione dei pezzi sulla scacchiera prima e dopo la mossa in questione. La domanda, allora, sarà se sia possibile o meno passare dall'una all'altra situazione in una sola mossa. Un giocatore potrebbe dire, per esempio: «Io l'ho messo in scacco e lui ha arroccato. Può farlo?». Oppure: «Io ho buttato cuori e lui mi ha risposto con picche, nonostante avesse in mano dei cuori. Può farlo?». Perché le regole di un gioco astratto dicono soltanto quali configurazioni di gioco si possono raggiungere in un'unica mossa da altre configurazioni; e, in certi casi, quali sono le configurazioni vincenti. In termini più tecnici, possiamo parlare di una configurazione di gioco come di uno *stato del gioco* o, più semplicemente, di uno *stato*, e del processo per cui da uno stato se ne raggiunge un altro come di una *transizione di stato*. Con questa terminologia, possiamo chiamare le regole di un gioco astratto *regole per le transizioni di stato*.

Tutti i giochi interessanti hanno regole permissive per le transizioni di stato. Di solito, cioè, il giocatore può scegliere tra un certo numero di mosse diverse, quando è il suo turno, tranne naturalmente in situazioni relativamente rare in cui la mossa è obbligata. Se non fosse così, il gioco in quanto tale sarebbe privo d'interesse; tutto il suo andamento, compreso il risultato finale, sarebbe determinato ancor prima dell'inizio. Resta tuttavia vero che in certi casi il risultato, benché predeterminato, non è noto al giocatore, e che questi può desiderare conoscerlo.

[2] Ci sono, naturalmente, molti giochi per cui non si è mai potuto dimostrare che le regole siano complete e coerenti. Quando, giocando a questi giochi, sorge un conflitto a causa dell'incompletezza delle regole, esso viene di solito risolto modificando le regole. Dopo un po', queste regole modificate vengono considerate e usate come regole «classiche» (devo questa osservazione a Oliver Selfridge).

Uno potrebbe voler sapere, mettiamo, che ora sarà 22 ore dopo le 9 del mattino. Per scoprirlo, dovrà giocare a una semplice versione del gioco matematico chiamato «aritmetica modulare». Un modo di esprimere la regola fondamentale di tale gioco è dire che *x mod z* rappresenta il resto della divisione di *x* per *z*. Nel nostro particolare esempio, il giocatore vuole sapere cosa segnerebbe l'orologio 22 ore dopo le 9, cioè alle 31. Il suo problema è quello di calcolare *31 mod 12* (ottenendo 7).

Ma proviamo a fare veramente un gioco su questo problema. Il piano di gioco consiste in 12 portacenere, inizialmente vuoti, allineati sulla tavola come in Figura 2.1. In più, si ha a disposizione un gran numero di ciottoli. Il giocatore inizierà facendo un mucchietto con un numero di ciottoli pari all'ora di partenza (9, nel nostro esempio); quindi aggiungerà al mucchio un numero di ciottoli pari al «numero di ore a partire da adesso» che ha in mente (22, nel nostro esempio). A questo punto, deve prendere un ciottolo dal mucchio e metterlo nel primo portacenere. Poi un altro nel secondo portacenere, e così via fino a quando ha esaurito i ciottoli, oppure ne ha messo uno nell'ultimo portacenere. In quest'ultimo caso, se il mucchietto non è esaurito, ricomincerà dal primo portacenere fino all'esaurimento dei ciottoli. A questo punto entra in gioco l'ultima regola: se sono rimasti dei portacenere vuoti, la risposta al problema è il numero di portacenere non vuoti; se tutti i portacenere sono vuoti, la risposta è 12; ma se in tutti i portacenere c'è almeno un ciottolo, il giocatore dovrà toglierne uno da ogni portacenere, e poi applicare di nuovo l'ultima regola.

Naturalmente, questo non è che un complicato gioco equivalente alla somma di due numeri e alla divisione del risultato per 12 con sottrazioni successive. Le regole del gioco non sono permissive; non permettono al giocatore di scegliere ogni mossa tra diverse possibilità. Al contrario, le regole stabiliscono

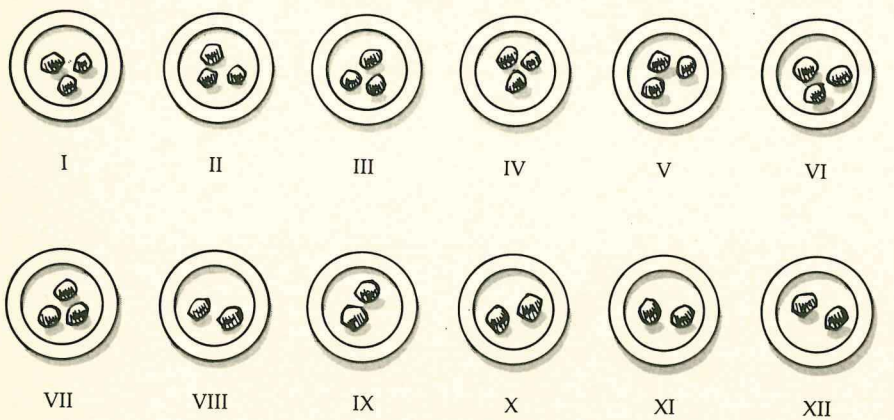


Figura 2.1. Il gioco dei portacenere per sapere l'ora

esattamente cosa il giocatore deve fare ad ogni passo. Un simile insieme di regole — vale a dire, un insieme di regole che dicono precisamente come comportarsi in ogni situazione — si chiama *procedura efficace*. Il concetto di procedura efficace, o «algoritmo» come viene anche chiamata, è uno dei più importanti della matematica moderna. Non solo la matematica è in gran parte volta a trovare procedure efficaci per fare molti tipi di operazioni utili, ad esempio lunghe divisioni, ma esistono anche problemi matematici di profondo significato, che hanno a che fare con la natura fondamentale della matematica stessa, che possono essere posti e affrontati in termini di ricerca di procedure efficaci e di discussione della loro validità.

La definizione di procedura efficace data sopra è ingannevolmente semplice. L'inganno sta nelle parole «dicono precisamente come comportarsi». Un giocatore che decida di risolvere il suo problema sulle ore seguendo le regole che ho appena enunciato, dovrà prima capire queste regole. Deve sapere cosa significa fare un mucchietto di ciottoli, che cos'è un portacenere, come decidere se un portacenere è vuoto (potrebbe contenere della cenere, ma nessun ciottolo), e così via. Deve, in altre parole, essere capace non solo di leggere le regole, ma anche di interpretarle esattamente nel modo in cui io volevo fossero interpretate. E se le regole devono dire «precisamente» come comportarsi, allora devono essere espresse in un linguaggio capace di produrre affermazioni precise. I libri di ricette, per esempio, sono procedure efficaci? Sicuramente, essi cercano di dire al cuoco cosa fare in ogni momento. Ma poi generalmente, anzi molto spesso, vi appaiono espressioni come «aggiungere un pizzico di paprika», «mescolare fino a raggiungere una certa consistenza», «condire a piacere». Chiunque sarebbe d'accordo che tali istruzioni sono ben poco precise. Tuttavia, possiamo immaginare una scuola di cucina che istruisca i suoi allievi a un tale livello di professionalità e di buon gusto che anche questo tipo di istruzioni, per vaghe che possano sembrare all'uomo della strada, abbiano per loro un significato preciso ed uniforme. Le ricette di quella scuola sembrerebbero allora essere procedure efficaci per i suoi diplomati, anche se non necessariamente per gli altri comuni mortali.

La mia descrizione di una procedura efficace come di un insieme di regole che dicono precisamente a un giocatore come comportarsi a ogni mossa sembra, dunque, essere difettosa, almeno nel senso che non regge in ogni circostanza. Dato un insieme di regole, ad esempio per cuocere un torta nel forno, sembra non esserci un criterio assoluto per determinare se tale insieme sia o meno una procedura efficace. Questa difficoltà non ci sarebbe, o almeno non per i libri di ricette, se fossero soddisfatte due condizioni: primo, l'esistenza di un linguaggio in cui fosse possibile enunciare regole di cucina precise e mai ambigue; secondo, un identico atteggiamento di tutte le persone di fronte a qualunque cosa abbia a che fare con la cucina. Tali condizioni non sono indipendenti una dall'altra, perché se tutte le persone pensassero allo stesso modo, finirebbero per interpretare allo stesso modo il linguaggio delle ricette. Ma neanche le più rigide scuole di cucina pretendono che i loro studenti diventino in tutto e per tutto uguali ai maestri. Sperano solo che i loro diplomati abbiano imparato a *imitare* i grandi chef nell'interpretazione delle ricette.

Questa prova di modestia da parte degli istituti di alta cucina può servire da esempio per noi, in quanto ci indica come procedere nel nostro discorso. Per conferire al concetto di «procedura efficace» uno status autonomo, abbiamo bisogno di un linguaggio in cui esprimere, senza ambiguità di alcun tipo, cosa un giocatore debba fare in ogni momento. Ma a quanto pare è possibile formulare procedure efficaci in molti linguaggi che, se non vincolati da regole espressamente costruite, sono intrinsecamente ambigui.

Il problema sarebbe risolto se esistesse anche un solo linguaggio intrinsecamente privo di ambiguità, in cui poter scrivere tutte le procedure efficaci. Sarebbe anche sufficiente usare quel linguaggio non già per scrivere le procedure efficaci che vogliamo eseguire, ma anche solo delle regole per interpretare altri linguaggi in cui tali procedure possono di fatto essere state scritte. Infatti, se a un individuo che conosce solo un linguaggio venisse data una procedura scritta in un linguaggio che non conosce, insieme con delle regole che indicano precisamente come interpretare le affermazioni nella lingua «straniera», egli sarebbe in grado di imitare il comportamento di qualcuno che, conoscendo tale lingua, seguisse quelle procedure. Abbiamo perciò bisogno di un linguaggio assolutamente privo di ambiguità, in cui poter scrivere procedure efficaci e con cui poter formulare regole per l'interpretazione di affermazioni fatte in altre lingue. Tali insiemi di regole dovrebbero, a loro volta, essere procedure efficaci, vale a dire procedure per l'interpretazione di espressioni del linguaggio cui vengono applicate. Ma in che linguaggio devono essere scritte queste regole? Sembra che siamo entrati in un circolo vizioso. Se avessimo a disposizione un linguaggio così — e vedremo che è possibile costruirne uno — potremmo riconoscere esattamente, data una procedura scritta in un linguaggio non ambiguo, che cosa essa ci dice di fare: quello che fa l'imitatore di cui parlavamo sopra. Quindi, ogni procedura di questo tipo avrebbe un'interpretazione unica, indipendente dal linguaggio in cui è stata originariamente scritta.

Non avendone discusso, ho finora dato l'impressione che, quando parlo di «linguaggi», io intenda non soltanto linguaggi formali come quello dell'aritmetica, ma anche lingue naturali come l'italiano o l'inglese. E infatti ho formulato le regole del gioco delle ore in italiano, citando però anche l'aritmetica. Tuttavia, vogliamo che i linguaggi di cui abbiamo appena discusso abbiano regole di interpretazione prive di ambiguità. Sappiamo che le lingue naturali sono note per la loro ambiguità. Più avanti discuteremo di cosa significhi «capire» una lingua naturale in termini formali. Per il momento, limitiamoci a considerare i linguaggi formali.

Un linguaggio formale è (di nuovo!) un gioco. Torniamo per un attimo al gioco degli scacchi. Consiste in un insieme di pezzi, in una scacchiera con una particolare configurazione, in una specificazione della posizione iniziale dei pezzi sulla scacchiera e in un insieme di regole per le transizioni di stato che dicono al giocatore come può passare da uno stato del gioco al successivo. Abbiamo già osservato che queste regole sono (tranne che in poche situazioni speciali) permissive; dicono al giocatore quali mosse *può* fare, non quali *deve* fare. Il fatto che lo stato iniziale sia specificato dalle regole è una peculiarità

degli scacchi, non una proprietà generale di tutti i giochi. Nel poker, si inizia dando a ogni giocatore cinque carte, ma nessuna regola stabilisce quali carte debbano essere. Naturalmente, ogni gioco deve avere inizio in qualche modo. Possiamo anche considerare l'inizio del gioco come uno «stato-zero» e poi aggiungere, alle già esistenti regole per le transizioni di stato, delle regole di formazione (anziché trasformazione), a volte permissive come nel poker, altre volte obbligatorie come negli scacchi; le regole di formazione ci dicono come trasformare lo stato-zero in quello che normalmente consideriamo essere lo stato iniziale del gioco, cioè le carte distribuite o i pezzi disposti sulla scacchiera.

I «pezzi» di un linguaggio formale sono il suo *alfabeto*, cioè l'insieme dei simboli che si possono manipolare in quel linguaggio. Se vogliamo continuare l'analogia con gli scacchi, possiamo pensare alla carta su cui i simboli del linguaggio vengono scritti come alla «scacchiera», ma non è essenziale. Le regole di transizione di un linguaggio formale hanno per esso la stessa funzione che le regole di transizione di un gioco hanno per il gioco: dicono al giocatore come muovere da uno stato del gioco ad un altro. Ho detto prima che l'unica domanda ragionevole che può essere posta all'arbitro di un gioco è se una certa mossa proposta sia lecita o meno. Esattamente lo stesso vale per i linguaggi formali. Tuttavia per i linguaggi formali, sebbene in senso stretto sia vero che solo quel tipo di domanda è possibile, è anche vero che essa può apparire in due forme diverse: «È lecita la transizione proposta?» oppure «La configurazione di simboli che stiamo discutendo è un'espressione ammissibile del linguaggio?». «Posso arroccare mentre il mio re è sotto scacco?» è una domanda del primo tipo per il gioco degli scacchi. «A questa configurazione si può arrivare giocando correttamente?» sarebbe una domanda del secondo tipo. Per certi tipi di configurazione, la risposta è semplice. Se riguardasse, ad esempio, una situazione in cui ci sono otto pedoni bianchi, e i due alfieri bianchi occupano caselle dello stesso colore, risponderemmo «No». Analogamente, se trovassimo due re dello stesso colore sulla scacchiera, e così via. Ma i giocatori di scacchi in realtà non pongono domande come queste, perché ogni partita inizia sempre dalla configurazione iniziale canonica, o al massimo ricomincia da una posizione raggiunta in una partita precedente, poi sospesa.

Molti linguaggi formali differiscono dagli scacchi, in questo. L'algebra che si studia alle scuole medie — non scenderò qui in dettagli sulle sue regole — ha, per esempio, delle regole di trasformazione per «mettere in evidenza» fattori comuni a più espressioni; per esempio, $ac + bc$ si trasforma in $(a + b)c$ seguendo una di queste regole. Ma, se si vuole poter applicare queste regole, le espressioni cui applicarle devono prima di tutto essere lecite in quel linguaggio, come «frasi grammaticalmente corrette». L'espressione $ac + bc +$ non è una frase corretta in termini algebrici, e non le si può applicare alcuna delle regole algebriche. Se si vuole giocare con l'algebra, dunque, bisogna prima disporre i pezzi sulla scacchiera in modo lecito. Bisogna sapere che un'espressione che inizia con una parentesi deve da qualche parte venire «chiusa» con una parentesi simile, che simboli di operazione come « + » vanno messi solo in mezzo a due espressioni, e così via.

Un linguaggio formale è un gioco. Questa non è soltanto una metafora, bensì un'affermazione che stabilisce una corrispondenza formale. Ma se tale affermazione è vera dovremmo essere capaci, quando parliamo di un linguaggio, di passare facilmente da un vocabolario di-tipo-gioco a uno, corrispondente, di-tipo-linguaggio e viceversa. Ed è esattamente quello che si può fare.

Descriverò ora un gioco in termini di scacchiera, pezzi, mosse e così via, per sviluppare alla fine i concetti linguistici corrispondenti. Il mio obiettivo è sviluppare un linguaggio molto preciso con un alfabeto molto ristretto; per di più, un linguaggio le cui regole di trasformazione possano essere scritte anch'esse usando soltanto quel piccolo alfabeto. Delineerò il progetto di una semplice macchina che incarna quelle regole di trasformazione. Tale macchina, naturalmente, sarà capace di giocare al gioco che ho descritto inizialmente. Scriverò poi le regole del gioco che ho progettato in un linguaggio il cui alfabeto corrisponda ai pezzi del gioco. È in quel linguaggio che il gioco consiste. Infine, indicherò come sia possibile progettare una seconda macchina che funga da interprete per il nuovo linguaggio. E vedremo dove tutto questo impianto ci può portare.

IL GIOCO

Attrezzatura necessaria:

Un rotolo di carta igienica.
Molte pietre bianche, cinque pietre nere e un normale dado da gioco.

Configurazione iniziale:

1) Srotolare parte del rotolo di carta igienica sul pavimento.
2) Metterci sopra delle pietre secondo il seguente schema:
   a) su un rettangolo scelto arbitrariamente, una pietra nera;
   b) sui rettangoli successivi, alla destra del precedente, tante pietre bianche quante volete, una per rettangolo di carta;
   c) andando ancora verso destra, una pietra nera, poi un rettangolo vuoto, poi ancora una pietra nera;
   d) sempre verso destra, un numero arbitrario di pietre bianche;
   e) a destra dell'ultima pietra bianca, una pietra nera;
   f) infine, un pietra nera (il «segnaposto») al di sopra del rettangolo che contiene la pietra bianca più a destra.
3) Girare il dado in modo che la faccia «1» sia rivolta verso l'alto.

(Un esempio di configurazione iniziale di questo tipo è mostrato in Fig. 2.2).

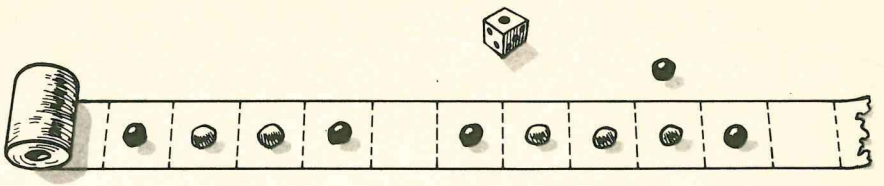


Figura 2.2. La configurazione iniziale del gioco

Le regole di trasformazione:

Il segnaposto viene spostato di una posizione a destra o a sinistra a ogni mossa. Però, prima di ogni mossa, la pietra che sta sotto il segnaposto viene spostata o sostituita secondo le regole elencate in Tabella 2.1. Ad ogni mossa, inoltre, è possibile che il dado debba essere girato in modo da mostrare un'altra faccia.

Tabella 2.1 *Le regole del gioco*

SE IL DADO SEGNA	E LA PIETRA SOTTO IL SEGNA POSTO È	ALLORA PORTARE IL DADO SUL	SOSTITUIRE LA PIETRA CON	E SPOSTARE IL SEGNAPOSTO A
1	nessuna	3	bianca	sinistra
1	nera	2	nessuna	sinistra
1	bianca	1	bianca	sinistra
2	nessuna	2	nessuna	sinistra
2	nera	3	nessuna	sinistra
2	bianca	5	nessuna	destra
3	nessuna	3	nessuna	sinistra
3	nera	4	nessuna	destra
3	bianca	5	nessuna	destra
4	nessuna	4	nessuna	destra
4	nera	1	nera	destra
4	bianca	6	bianca	sinistra
5	nessuna	5	nessuna	destra
5	nera	1	nera	destra
5	bianca	1	bianca	sinistra
6	nessuna	0	nessuna	destra
6	nera	0	nera	destra
6	bianca	3	bianca	sinistra

Le regole dicono: che cosa va sostituito alla pietra sotto il segnaposto (se una pietra bianca, o una nera, o nessuna); su quale faccia va girato il dado; in che direzione va mosso il segnaposto. Ci sono diciotto regole. Hanno tutte la stessa forma. Ciascuna di esse dice cosa fare quando il dado è orientato in un certo modo, e sotto il segnaposto c'è una certa pietra. Il giocatore deve trovare la regola corrispondente alla situazione di gioco che ha davanti, e poi applicarla. Ogni regola dice di fare tre cose:

1) Voltare il dado in modo che mostri il numero stabilito dalla regola.
2) Sostituire la pietra sotto il segnaposto con il tipo di pietra specificato (eventualmente, nessuna pietra).
3) Spostare il segnaposto di una posizione nella direzione indicata.

Seguendo una regola, si crea così una nuova situazione di gioco. Il giocatore applicherà di nuovo la regola opportuna, e così via. Quando la regola gli dirà di voltare il dado in modo che segni «0», vorrà dire che il gioco è finito.

Nella Tabella delle regole, le prime due colonne descrivono le condizioni in cui ciascuna regola è applicabile, mentre le restanti tre colonne dicono cosa fare in ciascun caso. Naturalmente, proprio come le sardine norvegesi nella Seconda Guerra Mondiale non erano una cosa da mangiare, bensì da vendere e comprare, così questo non è un gioco da giocare, ma piuttosto un gioco di cui parlare. E per poterne parlare più facilmente, cambiamo notazioni: invece di «pietra nera», «pietra bianca» e «nessuna pietra», useremo «X», «1» e «0», rispettivamente.

Una configurazione iniziale è quindi

...000X11X0X11<u>1</u>X00...

dove i puntini significano «e così via» (ricordare che tutto il rotolo di carta igienica è inizialmente fatto di «0», non contiene cioè alcuna pietra). Il segnaposto serve soltanto per aiutare la memoria; può essere sostituito dal dito indice del giocatore (ho usato la sottolineatura per indicare la sua posizione). Se ora interpretiamo le sequenze di 1 come numeri («111» significa «3»), vediamo che le X servono da punteggiatura, o meglio da virgolette che contengono due numeri. Data la configurazione di partenza di cui sopra, il gioco terminerebbe con la configurazione

...00<u>0</u>11111X00...

Quest'ultima può essere interpretata come somma dei due numeri presenti inizialmente. Il gioco costituisce così (anche se non l'ho dimostrato qui) una macchina per fare addizioni. La Tabella 2.2 mostra l'intera sequenza delle mosse per il caso di cui sopra; la colonna di destra indica il numero che appare ogni volta sul dado.

Tabella 2.2 *Il gioco*

MOSSA	PIANO DI GIOCO	PROSSIMA REGOLA
1	X11X0X111X	1
2	X11X0X111X	1
3	X11X0X111X	1
4	X11X0X111X	1
5	X11X00111X	2
6	X11X00111X	2
7	X11000111X	3
8	X10000111X	4
9	X10000111X	5
10	X10000111X	5
11	X10000111X	5
12	X10000111X	1
13	X10001111X	3
14	X10001111X	3
15	X10001111X	3
16	X10001111X	3
17	X00001111X	5
18	X00001111X	5
19	X00001111X	5
20	X00001111X	5
21	X00001111X	1
22	X00011111X	3
23	X00011111X	3
24	X00011111X	3
25	X00011111X	3
26	000011111X	4
27	000011111X	4
28	000011111X	4
29	000011111X	4
30	000011111X	6

Come esempio di applicazione delle regole, consideriamo la transizione dalla mossa 6 alla 7 nella Tabella 2.2. La riga 6 è

X11X00111X

con il dado sul 2. La regola da applicare deve dunque dire «Se il dado è sul 2 e sotto il segnaposto c'è una pietra del tipo X, cioè una pietra nera, allora fai questo e quest'altro». C'è una e una sola regola della Tabella 2.1 che ha questi requisiti: la quinta, che dice di:

1) Girare il dado sul 3.
2) Sostituire la pietra nera sotto il segnaposto con nessuna pietra; vale a dire, toglierla dal gioco.
3) Spostare il segnaposto di una posizione a sinistra.

Se si segue questa regola, la configurazione di gioco diventa quella mostrata nella riga 7,

X11000111X

e il dado mostra 3.

Ora provate a guardare l'evoluzione del gioco nel suo complesso, e in particolare notate come, in Tabella 2.2, il segnaposto oscilli avanti e indietro. Per aiutare l'intuizione, pensate che il gioco si svolga in un prato. Le pietre sono molto pesanti e il ragazzo che le sposta deve riposarsi ad ogni mossa. Possiamo allora riconoscere, a grandi linee, la sua strategia: prima cerca l'«1» più a destra tra quelli che formano il numero di sinistra (e lo trova alla mossa 7), e poi cerca un posto in cui metterlo. Per far ciò, deve trovare l'«1» più a sinistra tra quelli del numero di destra (cosa che fa alla mossa 11) e sostituire con un «1» il suo vicino di sinistra. Continuerà in questo modo fino a quando incontrerà la «X» che si trova più a sinistra e indica il limite del piano di gioco; a questo punto, saprà di aver finito. Naturalmente, può darsi che egli non si accorga della strategia che sta seguendo. Porta con sé le regole, e le consulta prima di ogni mossa.

È facile vedere come sarebbe possibile costruire una macchina per fare questo lavoro. Sostituiamo il rotolo di carta igienica con un comune nastro magnetico, e il giocatore con un comune registratore. Naturalmente, dobbiamo modificare lievemente le regole: ora è il nastro che si muove, anziché il segnaposto lungo il nastro. Perciò, laddove la regola diceva di spostare il segnaposto a sinistra, essa dirà ora di far muovere il nastro verso destra, e viceversa. I tre simboli di cui abbiamo bisogno possono essere rappresentati da tre suoni facilmente distinguibili. Installeremo poi sei relè nel registratore, per conservare l'informazione che prima era rappresentata dal dado (in realtà basterebbero tre relè, ma è più semplice usarne di più che addentrarsi in spiegazioni complicate). Supponiamo che «X» sia rappresentato da una nota molto

acuta, «1» da una media, «0» da una molto bassa. Se abbiamo registrato il nastro in modo corrispondente alla configurazione iniziale della Tabella 2.2, possiamo iniziare il calcolo spostando il nastro in modo tale che l'«1» che si trova più a destra (la nota media più a destra) si trovi sotto la testina del registratore (che può sia leggere che registrare), mentre il relè numero 1 è chiuso — visto che il dado segnava 1 — mentre tutti gli altri sono aperti. I circuiti della nostra macchina saranno congegnati in modo tale che:

> Se il relè 1 è chiuso e viene letta una nota alta (X), allora questa viene eliminata registrandoci sopra una nota bassa (0), il nastro viene spostato di una posizione a destra (il che equivale a spostare la testina, cioè il segnaposto, a sinistra), il relè 1 viene spento (aperto) mentre il relè 2 viene acceso (chiuso).

Questa è una fedele traduzione della seconda riga della nostra tabella di regole. Analogamente, ogni altra regola verrà tradotta nel «linguaggio» del registratore. A questo punto, avremo una macchina (incredibilmente goffa!) per fare le addizioni.

La cosa più importante da notare, in questa macchina, è il fatto che essa è completamente definita dalle regole del gioco, tradotte nel linguaggio del registratore. Se si assegnano quelle regole, è una macchina per fare addizioni; se le regole fossero diverse, sarebbe un altro tipo di macchina.

Si noti anche che abbiamo avuto bisogno di tre soli simboli distinguibili sul nastro magnetico. Li abbiamo indicati con «X», «1» e «0». Vedremo adesso che è possibile scrivere anche tutte le regole usando soltanto questi tre simboli!

Definiamo lo «stato» della macchina come il numero che appare sul dado, cioè il numero del relè che in quel momento è acceso (chiuso). Scriviamo ora le cinque parti delle nostre regole nel seguente ordine:

> stato attuale, simbolo sotto la testina, stato successivo, simbolo da registrare, direzione di moto del nastro

e adottiamo il seguente codice:

1) «X» è un segno di punteggiatura che indica l'inizio e la fine di un pezzo di informazione. È, in altre parole, una sorta di parentesi.
2) Quando vogliamo indicare un numero, scriveremo il corrispondente numero di «1».
3) Parlando di «Direzione di moto del nastro, «0» significa «a sinistra» e «1» significa a «destra».

Usando questo codice, possiamo ora riformulare le regole della Tabella 2.1 nella forma mostrata in Tabella 2.3.

Abbiamo scritto le regole su righe diverse, separando le varie colonne con degli spazi bianchi. Ma anche se eliminassimo questi spazi bianchi, e collegas-

simo le righe in una sola sequenza di «X», «1» e «0», non potrebbe risultarne alcuna confusione, nel senso che sarebbe ancora possibile interpretare la sequenza in modo univoco e privo di ambiguità. Quella sequenza costituirebbe allora una descrizione completa della nostra macchina per fare le addizioni!

Abbiamo così elaborato una notazione in cui possiamo descrivere una macchina. Il suo alfabeto consiste nei tre simboli «X», «0» e «1». Naturalmente, messaggi scritti usando questa simbologia rimarrebbero privi di significato, a meno di non avere un sistema per interpretarli. A questo scopo, sarebbe sufficiente descrivere una macchina — o, meglio ancora, costruirne una — che sia in grado di ricevere come dati d'ingresso i due blocchi di informazione determinanti: primo, la descrizione della nostra macchina per le addizioni, espressa in un codice opportuno; secondo, una configurazione iniziale di X, 0 e 1 su cui la macchina appena descritta deve operare. Queste due cosiddette *stringhe* (traduzione in «computerese» del termine inglese *string*, letteralmente «corda», NdT) di simboli potrebbero, naturalmente, essere messe sullo stesso nastro. Una volta in possesso di una macchina simile, abbiamo il diritto di chiamare il nostro sistema di notazioni «linguaggio», poiché a quel punto abbiamo un'incarnazione delle sue regole di trasformazione.

Tabella 2.3 *Le regole del gioco*

NUMERO DELLA REGOLA	SIMBOLO SOTTO LA TESTINA	PROSSIMO NUMERO DI REGOLA	SIMBOLO DA SCRIVERE	DIREZIONE DI MOTO DEL NASTRO
X1X	X0X	X111X	X1X	X1X
X1X	XXX	X11X	X0X	X1X
X1X	X1X	X1X	X1X	X1X
X11X	X0X	X11X	X0X	X1X
X11X	XXX	X111X	X0X	X1X
X11X	X1X	X11111X	X0X	X0X
X111X	X0X	X111X	X0X	X1X
X111X	XXX	X1111X	X0X	X0X
X111X	X1X	X11111X	X0X	X0X
X1111X	X0X	X1111X	X0X	X0X
X1111X	XXX	X1X	XXX	X0X
X1111X	X1X	X111111X	X1X	X1X
X11111X	X0X	X11111X	X0X	X0X
X11111X	XXX	X1X	XXX	X0X
X11111X	X1X	X1X	X1X	X1X
X111111X	X0X	XX	X0X	X0X
X111111X	XXX	XX	XXX	X0X
X111111X	X1X	X111X	X1X	X1X

In quello che fu uno dei più grandi trionfi dell'intelligenza umana, il matematico inglese Alan Turing dimostrò nel 1936 che una macchina simile poteva essere costruita, e mostrò anche come costruirla [3]. In realtà, egli dimostrò molto di più; ma di questo riparleremo più avanti. Non mi è possibile qui descrivere nei dettagli una macchina costruita secondo i principi enunciati da Turing, ma devo comunque farvi un accenno.

Immaginate di nuovo un registratore a nastro come quello che abbiamo usato nella nostra macchina per addizioni. Anche ora, supponiamo che abbia un insieme di relè capaci di rappresentare i suoi diversi stati. Questa volta, essendo la macchina molto più complicata di quella per le addizioni, ci sono molti più stati da rappresentare, e quindi molti più relè; ma questo è un dettaglio non essenziale. Il nastro usato dalla nuova macchina ha la seguente struttura; leggendo da destra verso sinistra:

un tratto che contiene la descrizione della nostra macchina per addizioni, nella notazione che abbiamo elaborato sopra;

un tratto che contiene i dati su cui quella macchina deve lavorare, per esempio «X11X0X111X»;

un tratto che può immagazzinare, in ogni momento, lo stato della «macchina per addizioni» in quel momento;

un tratto arbitrariamente lungo di nastro «vuoto», cioè un tratto contenente soltanto degli «0».

L'informazione contenuta nel nastro ha quindi la seguente struttura:

nastro vergine	stato attuale	simbolo attuale	dati	descrizione della macchina

Per vedere come la macchina funzioni, si deve solo fare finta di aver ricevuto una descrizione della macchina per fare le somme, gli insiemi di dati necessari e alcune convenzioni, come quella per cui la macchina deve sempre partire dallo stato 1, o quella per cui la «X» all'estrema destra del suo blocco di dati è un simbolo che indica l'inizio dei dati alla sua sinistra. In più, tornerà utile della carta su cui prendere appunti. È proprio così che fa la macchina: fa andare avanti e indietro il nastro, legge dati, segna appunti sul nastro di scorta e consulta il tratto che contiene la descrizione, per sapere come procedere. In questo modo, molto lentamente ma passo passo, e in modo estrema-

[3] A.M. Turing, *On Computable Numbers, with an Application to the* Entscheidungsproblem («Sui numeri calcolabili, con un'applicazione all'*Entscheidungsproblem*»), Proc. London Math. Soc. Ser. 2-42 (17 novembre 1936), pp. 230-265.

mente fedele, la nuova macchina imita la nostra iniziale macchina per fare addizioni, cioè la macchina che ha, descritta, sul suo nastro. Questa nuova macchina, dunque, incarna davvero le regole che spiegano come interpretare le stringhe di X, 1 e 0 che le abbiamo dato.

Una macchina del tipo appena discusso, cioè una macchina che fa muovere un nastro magnetico avanti e indietro, leggendo e cambiando i simboli su un pezzetto di nastro per volta, passando da uno stato ad un altro e così via, viene chiamata oggi «macchina di Turing». Una macchina simile è descritta completamente quando, per ogni stato che può raggiungere e per ogni simbolo che può trovarsi sotto la sua testina, la descrizione indica quale simbolo deve scrivere, in quale stato deve passare e in che direzione deve far muovere il nastro. Possiamo stabilire, come convenzione, che ogni macchina di questo tipo parta sempre dal suo stato iniziale, che chiameremo «stato 1», e con il nastro posizionato in modo tale che il simbolo più a destra sia sotto la testina di lettura/scrittura. Qui abbiamo deciso di usare un linguaggio il cui alfabeto consiste nei tre simboli che abbiamo adoperato per descrivere la nostra macchina per addizioni. Se fossimo stati più generosi, vale a dire se ci fossimo concessi l'uso di un alfabeto più esteso, la nostra macchina avrebbe potuto essere più semplice, nel senso di avere un minor numero di possibili stati. D'altra parte, le regole sarebbero state più complesse. Ci sono, quindi, molti modi possibili per realizzare la nostra macchina che fa le somme; almeno uno per ogni possibile scelta dell'alfabeto. Il numero minimo di caratteri per un alfabeto che si potrebbe usare è due; ma in tal caso una macchina come la nostra dovrebbe avere un gran numero di stati su cui operare. Il numero minimo di stati che una macchina per le addizioni potrebbe avere è due; ma dovrebbe allora operare su un alfabeto molto esteso.

C'è una differenza importante tra le due macchine che abbiamo discusso: la nostra macchina per le addizioni è una macchina con un particolare obiettivo. Sa sommare due numeri qualsiasi, ma non sa fare nient'altro. La seconda macchina ha bisogno, come *input*, di una descrizione codificata di un'altra macchina e di un blocco di dati su cui la macchina descritta deve operare. In effetti, la descrizione che la seconda macchina riceve è un *programma* che la trasforma nella macchina che deve imitare. E sorge in modo naturale la domanda: «Quali macchine possono venire imitate in questo modo?».

Ho spiegato cosa intendo parlando di un linguaggio formale, e cioè: un alfabeto; un insieme di regole di formazione che determinano il formato delle stringhe di simboli, costruite basandosi su quell'alfabeto, che costituiscono espressioni lecite in quel linguaggio; e un insieme di regole di trasformazione per quelle espressioni. Ho anche detto che è possibile costruire macchine che incarnano tali regole di trasformazione, e possono perciò eseguire procedure rappresentate nei linguaggi corrispondenti. Al di là di questo, ho discusso una macchina che accetta descrizioni di altre macchine, ed è capace di imitarne il comportamento. Ci siamo così fatti un'idea di cosa si intenda, in questo contesto, per imitazione. Fate bene attenzione al fatto che non ho mai parlato di traduzione da un linguaggio a un altro. La nostra macchina imitatrice non traduce le regole di trasformazione che le abbiamo dato, cioè la descrizione codi-

ficata della macchina per fare addizione, nel suo o in alcun altro linguaggio. Essa consulta — diciamo di solito che «interpreta» — quelle regole di trasformazione ogni volta che deve decidere che cosa farebbe la macchina che sta imitando. Quindi, fa molte mosse per ogni mossa della macchina da imitare.

Il mio obiettivo, per il momento, è quello di dare una solida base al nostro concetto di procedura efficace. Vogliamo trovare un linguaggio in cui poter esprimere le procedure efficaci, almeno nel senso di poter descrivere in quel linguaggio tutti i nostri linguaggi procedurali, dando così un'interpretazione univoca alle nostre procedure. Ora sappiamo che le regole di trasformazione dei linguaggi possono essere incarnate in certe macchine. Mi resta dunque da dimostrare che è possibile trovare un unico alfabeto, e un linguaggio basato su quell'alfabeto, in cui davvero possiamo descrivere tutti i linguaggi in cui ci può capitare di scrivere procedure.

Possiamo progettare un linguaggio il cui alfabeto consiste solo in due simboli, mettiamo «0» e «1», e tramite il quale siamo in grado di descrivere qualsiasi macchina di Turing. Ora, abbiamo visto che un linguaggio consiste in qualcosa di più che non nel solo alfabeto, proprio come un gioco non consiste soltanto nei pezzi con cui viene giocato. Bisogna anche assegnare le sue regole di trasformazione. In questo momento, penso alle regole di trasformazione del linguaggio che ho in mente come a regole incarnate in una macchina di Turing, simile alla macchina imitatrice che abbiamo già discusso. Sto dicendo, quindi, che esiste una macchina di Turing che opera su un nastro contenente solo «1» e «0», e che è capace di imitare qualunque altra macchina di Turing, nessuna esclusa. Questa cosiddetta macchina di Turing *universale* è, come tutte le macchine di Turing, descrivibile in termini di insieme di «quintuple» del tipo che già conosciamo, e cioè

> (stato attuale, simbolo letto, stato successivo, simbolo scritto, direzione di moto del nastro)

e queste quintuple, a loro volta, possono essere scritte nel linguaggio che la macchina di Turing può accettare, secondo il suo progetto. Per conseguenza, la macchina di Turing universale è capace di accettare una descrizione di se stessa, e di imitare se stessa.

In realtà, si possono progettare molti linguaggi con lo stesso alfabeto a due simboli, vale a dire molte macchine di Turing universali che incarnano regole di trasformazione su stringhe di questi due numeri. E si può, naturalmente, allargare l'alfabeto e progettare molte macchine di Turing universali corrispondenti a ciascuno di questi allargamenti. Ma quello che ci interessa in questo momento è il principio, vale a dire il fatto che

> esiste una macchina di Turing U (in realtà, un'intera classe di macchine) il cui alfabeto consiste nei due simboli «0» e «1» e tale che, data qualunque procedura scritta in un linguaggio preciso e privo di ambiguità e data una macchina di Turing L che incarna le regole di trasformazione di quel linguaggio, la macchina di Turing U sa imitare la L nell'esecuzione di quella procedura.

Questa è una riformulazione, con parole diverse, del risultato davvero notevole che Turing annunciò nel suo brillante lavoro del 1936.

In matematica ci sono molte «dimostrazioni di esistenza». Ma c'è una grande differenza tra l'essere capace di dimostrare che qualcosa esiste e l'essere capace di costruirla. Turing dimostrò che una macchina di Turing universale esiste indicando come costruirla. Dobbiamo ricordare che Turing fece questo lavoro di importanza monumentale nel 1936, più o meno un decennio prima che i primi computer moderni fossero davvero costruiti. I computer moderni non assomigliano alla macchina descritta da Turing. Molti di essi hanno, per esempio, la capacità di manipolare molti nastri magnetici simultaneamente; cosa anche più importante, per lo più essi sono dotati di spazi enormi in cui immagazzinare le informazioni. Dal punto di vista funzionale, il «magazzino» di un computer moderno è come un insieme di relè, ciascuno dei quali può essere acceso (chiuso) o spento (aperto).

Un insieme di dieci di questi relè può assumere 1024 stati diversi. Non è eccezionale, per un computer moderno di medie dimensioni, avere più di un milione di questi componenti elementari di «memoria», e quindi essere capace di assumere $2^{1.000.000}$ stati. Questo è un numero così grande che va al di là dell'immaginazione (la Terra, per esempio, pesa molto meno di $2^{1.000}$ chili). Ciononostante, in linea di principio ogni computer moderno è una macchina di Turing. Per di più ogni computer moderno, tranne pochi super-specializzati, è una macchina di Turing *universale*. E questo, in pratica, significa che ogni computer può, almeno in linea di principio, imitare qualunque altro computer moderno.

Resta ancora una lacuna da colmare. Infatti, anche dato per scontato che veramente ogni computer possa fare tutto ciò che qualunque altro computer sa fare, resta da stabilire *che cosa* tutti questi computer possano fare; vale a dire, per quali procedure sia possibile costruire della macchine di Turing, che possono essere imitate da macchine di Turing universali e quindi dai computer moderni. Alan Turing ha dato una risposta anche a questo problema: è possibile costruire una macchina di Turing per realizzare qualsiasi procedimento che possa venir considerato in modo naturale una procedura efficace.

Questa tesi, spesso chiamata «tesi di Church» in omaggio al logico Alonzo Church, che la formulò in un contesto diverso da quello di Turing, non può essere dimostrata, in quanto contiene l'espressione «in modo naturale». In un certo senso, siamo bloccati in un circolo vizioso logico; qualunque procedimento possiamo descrivere in termini di macchina di Turing è una procedura efficace, e viceversa. Ciò che, tuttavia, rafforza almeno dal punto di vista intuitivo il nostro concetto, è il fatto che parecchie formulazioni dell'idea di «calcolabilità efficace», radicalmente diverse fra loro e derivate in modo indipendente, si sono rivelate equivalenti alla calcolabilità secondo il formalismo di Turing, e quindi equivalenti fra loro. Come osserva M. Minsky: «La dimostrazione dell'equivalenza di due o più definizioni ha conseguenze vincolanti, quando tali definizioni provengano da esperienze e motivazioni diverse» [4].

[4] M. Minsky, *Computation, Finite and Infinite Machines* («Calcolo, macchine finite e infinite»), Prentice-Hall, Englewood Cliffs N.J., 1967, p. 111.

Ma anche se dobbiamo affidarci all'intuizione per decidere cosa può essere chiamato «in modo naturale» procedura efficace, abbiamo ora delle solide basi per riconoscere in modo preciso e privo di ambiguità che cosa una procedura efficace ci dice di fare. Almeno in linea di principio, sappiamo codificare l'alfabeto del linguaggio in cui la procedura è scritta usando soltanto i due simboli «0» e «1». Possiamo quindi trascrivere le regole che costituiscono la procedura nella nuova notazione. Infine, possiamo trasmettere — sotto forma di quintuple opportunamente codificate — le regole di trasformazione del linguaggio originale a una macchina universale di Turing che opera sull'alfabeto 0,1. La procedura da cui siamo partiti ci dice di fare quello che la macchina di Turing universale che abbiamo istruito fa quando imita la «macchina» che le abbiamo descritto. Se siamo in grado di capire come funziona una macchina di Turing (e abbiamo visto che non ci vuole molto a capirlo) e se abbiamo una descrizione della macchina di Turing universale cui ci siamo riferiti, allora sapremo nei dettagli che cosa la procedura di partenza ci dice di fare.

Una simile forma di conoscenza è assai debole. Non ci sogneremmo mai di dire che conosciamo una città — per non parlare del capirla — soltanto perché ne abbiamo una mappa dettagliata. E, a parte queste considerazioni, se comprendiamo il linguaggio in cui una procedura è scritta abbastanza da saperne le regole di trasformazione, probabilmente capiremo anche che cosa delle regole espresse in quel linguaggio ci dicono di fare.

Ma queste obiezioni, per quanto valide, vanno fuori strada. La tesi di Turing ci dice che è possibile realizzare, in forma di programma per un computer, qualsiasi procedura che possa essere «in modo naturale» chiamata procedura efficace. Perciò, ogni volta che crediamo di capire un fenomeno, nel senso di conoscerne le regole di comportamento, dovremmo essere capaci di esprimere la nostra comprensione sotto forma di programma per computer. Turing dimostrò che tutti i computer (a parte quelli più strettamente specializzati, di cui qui non ci interessiamo) sono equivalenti fra loro, cioè sono universali. Quindi, se un computer ben funzionante dal punto di vista tecnico non si comporta esattamente come credevamo di averlo programmato a fare, non si può attribuire la colpa a una particolarità di *quel* computer. In realtà, lo sbaglio deve stare nella nostra disattenzione nel trascrivere le regole di comportamento, che crediamo di capire, nel linguaggio formale che il computer richiede; oppure, nella formulazione di ciò che avevamo in mente quando abbiamo creduto di capire; oppure, nel fatto che la nostra comprensione è stata incompleta. Quest'ultima è quasi sempre la vera ragione. È questo un argomento di cui riparlerò diffusamente più avanti. Per ora, ci serve solo osservare che l'incompletezza della nostra comprensione può assumere due forme.

In primo luogo, anche se la nostra teoria è corretta nel suo complesso, essa può contenere un errore in qualche dettaglio. Da qualche parte, per esempio, affermiamo erroneamente che se questo e quest'altro sono veri, allora ne consegue che... I nostri processi mentali, compiaciuti magari dell'eleganza del ragionamento che stiamo conducendo, ci permettono spesso di passar sopra a errori di quel tipo senza la minima preoccupazione. Il computer, però, è davvero spietato in queste cose. Esso segue rigorosamente la logica che gli

abbiamo dato. E quella logica può condurre a conseguenze molto diverse da quelle cui ci portano i nostri processi mentali, contaminati dal desiderio di ottenere certi risultati. Anzi, una delle migliori ragioni per l'uso dei computer è la possibilità di mettere in evidenza lacune nel nostro modo di pensare. I computer sono dei critici impietosi, da questo punto di vista.

In secondo luogo, il problema può essere che, anche se è vero che abbiamo capito la teoria, non siamo ancora in grado di formalizzare la nostra comprensione. Possiamo, per esempio, essere capaci di prevedere con grande sicurezza che cosa un animale farà in molte situazioni diverse. Ma la nostra capacità preditiva, per quanto estesa e affidabile, si basa magari su intuizioni che siamo incapaci di esplicitare in modo adeguato. Ciononostante, possiamo essere tentati di inserire ugualmente le nostre idee in uno schema formale. Un programma per computer basato su un sistema formale derivato in questo modo sicuramente non funzionerà. Il problema, a quel punto, non è più soltanto che la teoria rappresentata contiene degli errori di dettaglio, bensì che essa è fondamentalmente sbagliata in ciò che afferma sul suo argomento. Non sempre è chiaro dove stia il difetto, quando il computer che abbiamo programmato non funziona. Di solito, si è enormemente motivati a credere che le nostre teorie sono esatte nel loro complesso e che, se non funzionano bene, ci deve essere qualche dettaglio sbagliato cui è facile porre rimedio. Avrò qualcos'altro da dire su questo argomento più avanti.

Poco fa abbiamo introdotto l'idea dell'universalità dei computer. Ci dobbiamo ora chiedere se tale idea significhi che i computer possono «fare qualsiasi cosa». La domanda in realtà è: «È possibile descrivere qualunque cosa possiamo voler fare in termini di una procedura efficace?» La risposta è «No».

Per prima cosa, ci sono domande che possiamo fare e alle quali (lo si può dimostrare) non si può dare risposta usando alcuna procedura efficace. Potremmo, per esempio, essere interessati a sapere se una certa macchina che abbiamo progettato, per esempio la nostra macchina per addizioni, si fermerà a un certo punto dell'esecuzione del suo compito. Sarebbe opportuno avere una macchina di riferimento che, data una qualsiasi altra macchina con i suoi dati di ingresso, ci dica se tale macchina si fermerà durante l'elaborazione di quei dati. Ebbene, una macchina di questo tipo non può essere costruita. Questo problema, e con esso molti altri altrettanto «irresolubili», impone perciò dei limiti a ciò che i computer possono fare. Naturalmente queste sono limitazioni logiche, che vincolano non soltanto i calcolatori elettronici ma qualunque calcolatore, sia esso umano o meccanico. Bisogna anche dire che tutto l'insieme di problemi irresolubili non è terribilmente interessante dal punto di vista pratico; si tratta sempre di problemi molto generali. Se volessimo sapere se un particolare calcolo terminerà o meno, probabilmente potremmo progettare una procedura per scoprirlo. Ciò che è impossibile, è avere una macchina — o, che è lo stesso, una procedura efficace — che sappia scoprirlo per qualunque procedura *in modo generale*.

In secondo luogo, una procedura efficace può essere capace di fare certi calcoli in linea di principio, ma può richiedere un tempo così lungo da svuotarla di ogni utilità pratica. Pensate al gioco degli scacchi, per esempio. Data

la regola secondo cui il gioco termina se si ritorna per la terza volta alla stessa configurazione sulla scacchiera, gli scacchi sono certamente un gioco con un numero finito di mosse. È perciò possibile, in linea di principio, scrivere una procedura che generi un elenco di tutte le partite che potrebbero essere giocate. Ma un calcolo simile richiederebbe miliardi di anni sul più veloce computer che possiamo immaginare. È, quindi, un esempio di procedura priva di valore pratico. Per la verità, finora abbiamo discusso di procedura facendo finta che il tempo richiesto per la loro esecuzione fosse un aspetto irrilevante. Un simile atteggiamento è giustificato finché si parla di giochi astratti. In pratica, naturalmente, il tempo è importante. Dobbiamo in particolare notare che, quando un computer ne imita un altro, dovrà compiere molti passi, che richiedono tempo, per ogni passo del computer imitato. Se non fosse così, cercheremmo solo di costruire la macchina di Turing universale più economica possibile: questa, essendo capace di imitare ognuna delle macchine più costose, le eliminerebbe dal mercato.

In terzo luogo, potremmo scrivere una procedura realizzabile da una macchina di Turing, quindi una procedura efficace, le cui regole però non ne comprendono una che stabilisca quando fermarsi. Per esempio, la procedura «Partendo da zero, aggiungi uno e, se la somma è maggiore di zero, aggiungi di nuovo uno, e così via» ovviamente non finisce mai. Potremmo scrivere «se la somma è minore di zero, fermati, altrimenti aggiungi ancora uno» al posto di «se la somma è maggiore di zero, aggiungi ancora uno»; così, la procedura avrebbe una regola che dice quando fermarsi. Però, un calcolo che segua quella procedura non incontrerebbe mai tale regola; vale a dire, la macchina di Turing corrispondente non si verrebbe mai a trovare nello stato in cui «la somma è minore di zero». La procedura perciò è, in un certo senso, difettosa. Non è sempre facile (per dirla con un eufemismo) riconoscere se una vera procedura, scritta per un vero computer, sia o meno immune da difetti di questo tipo.

Infine, arriviamo all'aspetto più problematico del problema di cosa i computer possano e non possano fare. Ho detto più volte che una procedura efficace è un insieme di regole che ci dice, in un linguaggio preciso e privo di ambiguità, cosa fare in ogni momento. Ho sostenuto che un linguaggio è preciso e privo di ambiguità soltanto se il suo alfabeto e le sue regole di trasformazione possono essere essi stessi resi espliciti in termini precisi e non ambigui. E ho ripetuto la tesi di Church (e di Turing) secondo cui a ogni simile esplicitazione di un linguaggio corrisponde una macchina di Turing, che può essere imitata da una macchina di Turing universale. Inoltre, ho affermato che sostanzialmente tutti i computer moderni sono macchine di Turing universali. Lasciando ora da parte tutto ciò che riguarda questioni formalmente indecidibili, procedure interminabili e procedure difettose, ci troviamo di fronte alla domanda inevitabile: «È possibile ridurre tutti i processi adottati dall'uomo a procedure efficaci, e quindi farli eseguire dalle macchine?».

Abbiamo visto che l'idea stessa di procedura efficace è inestricabilmente legata all'idea di linguaggio. Non è strano che io abbia potuto discutere così a lungo del linguaggio senza mai citare il *significato*? La ragione per cui ho potuto evitare di affrontare il concetto di significato sta nel fatto che ho di-

scusso soltanto di linguaggi formali o, come ho detto, di giochi astratti. Non che il significato non abbia un ruolo in questi giochi di linguaggio. Ce l'ha. Ma questo ruolo è completamente implicito nelle regole di trasformazione del linguaggio. Ricordate che in algebra possiamo trasformare $ac + bc$ in $(a + b)c$: abbiamo il diritto di dire che le due espressioni significano la stessa cosa o, per dirla altrimenti, che la regola di trasformazione che abbiamo usato conserva il «valore» dell'espressione originale. Ancora in altre parole, se dovessimo sostituire *a, b,* e *c* con dei numeri, le due espressioni produrrebbero lo stesso risultato quando si eseguano le operazioni aritmetiche indicate. (Quest'ultima, sia detto per inciso, non è una proprietà di tutte le algebre. L'algebra elementare è stata intenzionalmente costruita in modo tale che le sue regole di trasformazione siano coerenti con quelle del linguaggio formale che chiamiamo aritmetica). È proprietà dei linguaggi formali, anzi è loro essenza, il fatto che tutte le loro regole di trasformazione sono puramente sintattiche, cioè descrivono i possibili modi di riordinare stringhe di simboli del linguaggio, compresa la sostituzione di simboli e l'introduzione di simboli nuovi — per esempio «)» e «(» — indipendentemente da qualsiasi significato tali simboli possano avere al di fuori dello schema del linguaggio stesso. Per esempio, è possibile riempire pagine di trasformazioni algebriche, seguendo ciecamente le regole dell'algebra, senza dover necessariamente sapere che è possibile sostituire le lettere minuscole (ma non le parentesi!) con dei numeri; in altre parole, senza mai interpretare i simboli che si stanno usando. Lo stesso non vale per il linguaggio naturale. Prendiamo la frase: «Non ho mai visto un uomo alto come Antonio». Essa può essere trasformata in «Antonio è l'uomo più alto che abbia mai visto», e il significato della frase originale è conservato. Ma se applichiamo la stessa regola di trasformazione a «Non ho mai visto un uomo alto come Paola», ottenendo «Paola è l'uomo più alto che abbia mai visto», non funziona più. La regola che abbiamo applicato non è puramente sintattica. Riguarda non soltanto la *forma* di una sequenza non interpretata di simboli, ma anche il loro *significato*.

Abbiamo visto che, fino a un certo livello del discorso, non c'è alcuna differenza sostanziale tra un linguaggio e una macchina che ne incarni le regole di trasformazione. Abbiamo anche osservato che, sebbene le leggi di cui le macchine astratte sono incarnazione non debbano essere necessariamente coerenti con le leggi dell'universo fisico, le leggi incarnate da macchine che interagiscono col mondo reale devono per forza di cose essere un sottoinsieme delle leggi che governano il mondo materiale. Se vogliamo continuare a identificare i linguaggi con le macchine anche quando discutiamo la lingua naturale, dobbiamo renderci conto che, qualunque siano le macchine che corrispondono ai linguaggi naturali, esse sono più simili alle macchine che trasformano energia e distribuiscono potenza meccanica che alle macchine astratte che abbiamo analizzato più sopra. Vale a dire, le loro leggi dovranno tenere conto del mondo reale. Anzi, esse saranno sottoposte a vincoli ancora più forti di quelli imposti ai semplici motori. Infatti, mentre le leggi dei motori sono soltanto un sottoinsieme delle leggi della fisica, le leggi di una macchina per il linguaggio naturale devono in qualche modo corrispondere alle realtà interiori, manifeste e

latenti, di chi parla quel linguaggio nel momento in cui lo parla. La lingua naturale è difficile in questo senso perché dobbiamo sapere, per esempio, per quali «valori» di *X* e *Y* è lecito applicare la regola di trasformazione che ci porta da «Non ho mai visto un *X* alto come *Y*» a «*Y* l'*X* più alto che abbia mai visto». Chiaramente, non è una regola applicabile in modo uniforme a qualunque sequenza di simboli, indipendentemente dalla loro interpretazione.

Questa difficoltà è anche più profonda di quanto possa apparire a prima vista. Perché non è nemmeno possibile definire il dominio di applicabilità di questa regola — e ce ne sono molte analoghe — usando, mettiamo, elenchi di sostantivi, pronomi e nomi propri maschili e femminili per correggere opportunamente la regola. Prendiamo una variazione dello stesso esempio che ho già usato, e cioè la frase «George è l'uomo più intelligente che abbia mai incontrato». Immaginate un romanzo giallo in cui fin dal primo capitolo risulti chiaro che il criminale da scroprire deve essere una persona che si cela sotto mentite spoglie. L'infallibile ispettore smaschera l'impostore, diciamo, nel capitolo 10. Lo scopo del capitolo 11 è spiegare, al lettore che non è stato capace di dedurre le conclusioni dell'ispettore dagli indizi sparsi nei capitoli precedenti, come abbia fatto l'ispettore a identificare il colpevole. Nel capitolo 11, dunque, l'ispettore ricorda che il signor Micavero aveva pronunciato la frase «George è l'uomo più intelligente che abbia mai incontrato» in un salotto letterario in cui si discuteva dell'opera di George Eliot. Il signor Micavero si era fatto invitare in quel salotto facendo credere alla padrona di casa di essere un'autorità in materia di letteratura inglese del diciannovesimo secolo. L'ispettore si era reso conto che chiunque sapesse qualcosa della letteratura inglese avrebbe saputo che George Eliot era lo pseudonimo usato da Mary Ann Evans, una donna. L'osservazione del signor Micavero era perciò «sbagliata dal punto di vista grammaticale», più o meno come l'espressione matematica x/y è sbagliata grammaticalmente quando $y = 0$. Quindi, il signor Micavero non poteva essere quello che sosteneva di essere.

Un buon romanzo giallo (o forse dovremmo dire uno «onesto») dà al lettore tutte le informazioni necessarie a scoprire la verità, per esempio chi è l'autore del delitto, prima di spiegare come il detective sia arrivato alle sue deduzioni. Il punto-chiave di un romanzo giallo spesso sta proprio nel fare uscire dall'ambiguità ciò che uno dei personaggi aveva detto all'inizio. E la possibilità stessa di riconoscere un'ambiguità, e quindi di sapere che essa richiede una spiegazione, insomma la possibilità di risolvere il mistero, dipende dalla conoscenza che il lettore ha del mondo reale, e dalla proprietà del linguaggio naturale per cui le sue regole si applicano a sequenze di simboli interpretate nel contesto del mondo reale. Una storia come quella che abbiamo appena discusso non può essere compresa in termini puramente formali. Ed è interessante il fatto che neppure il primo capitolo può venire tradotto in un'altra lingua se il traduttore non conosce, e non ha capito, l'errore fatale del signor Micavero; senza, cioè, una comprensione delle rivelazioni fatte nell'ultimo capitolo.

Ritorneremo più avanti all'esame del ruolo che il contesto ha nella comprensione del linguaggio naturale, da parte degli uomini o da parte delle macchine. Per ora, il nostro interesse resta limitato al problema più ristretto —

almeno secondo la nostra formulazione — della possibilità di convertire i processi decisionali umani in procedure efficaci, e quindi in processi calcolabili.

Ci sono, naturalmente, processi decisionali umani che possono essere descritti in modo chiaro e non ambiguo anche nel linguaggio naturale. Poco fa ho descritto dei giochi sia usando il linguaggio naturale che in termini di progetti di macchine, descritti a loro volta in lingua naturale. E in realtà non potremmo mai capire una macchina di Turing, o una procedura efficace tradotta in termini di macchina di Turing (vale a dire, un programma per una macchina di Turing universale), se non fossimo prima in grado di capire che cosa significa dire che una porzione di nastro magnetico è adiacente a un'altra, o cosa significa leggere un simbolo da (o scriverlo su) una porzione di nastro magnetico, o cosa significa spostare il nastro di una posizione a sinistra o a destra, e così via. Ciò che è straordinario è il numero incredibilmente limitato di cose che bisogna sapere per poter avere accesso, in linea di principio, a tutta la matematica. Nella vita di tutti i giorni ci scambiamo istruzioni, cioè ci descriviamo reciprocamente procedure, che, benché siano magari tecnicamente ambigue in quanto sono potenzialmente soggette a diverse interpretazioni, a tutti gli effetti pratici sono procedure efficaci. In ultima analisi, esse si basano su un vocabolario comune a gran parte della società umana, un vocabolario i cui elementi hanno sostanzialmente un'unica interpretazione, quando vengono usati in contesti fortemente convenzionali. La maggior parte dei colloqui professionali e tecnici si valgono quasi esclusivamente di vocabolari di questo tipo. Il problema della conversione di tali procedure in procedure efficaci nel senso tecnico della parola, cioè in programmi per le macchine di Turing, è fondamentalmente una questione di formalizzazione della base di conoscenza che sta dietro all'interpretazione convenzionale dei loro vocabolari. Quanto più «standardizzati» sono questi vocabolari, e quanto più ristretto è il contesto in cui vengono usati, tanto più probabilmente sarà possibile risolvere questo problema. Infatti, dopo tutto, se ogni simbolo di un certo insieme di simboli ha un'interpretazione sostanzialmente unica in un certo contesto, e se sequenze di simboli vengono trasformate soltanto secondo regole che nascono anch'esse da quel contesto, allora non vi sarà alcuna ragione formale per far sorgere il problema dell'interpretazione da dare a ciascun simbolo. Un linguaggio così vincolato è sostanzialmente un linguaggio formale. Le sue regole, quindi, sono potenzialmente realizzabili da una macchina di Turing.

Ma rimangono le molte decisioni che prendiamo nella vita di tutti i giorni, per le quali non sappiamo descrivere in modo chiaro un processo decisionale. Come decido qual è la prossima parola che scriverò? Forse la nostra incapacità in questo campo è dovuta al fatto che, finora, non siamo riusciti ad arrivare a una comprensione sufficiente del linguaggio umano, della mente, del cervello, della logica simbolica. In fondo, visto che tutti noi possiamo imparare a imitare le macchine di Turing universali, siamo noi stessi per definizione delle macchine di Turing universali. Per meglio dire, siamo *almeno* delle macchine di Turing universali. (Anche una macchina di Turing fisicamente costruita non è soltanto una macchina di Turing; può, per esempio, essere anche un fermacarte o un sostegno per un vaso di fiori).

Ci uniamo qui a Michael Polanyi nel dire che sappiamo più di quanto riusciamo a dire[5]. Ma così dicendo siamo arrivati a chiudere il circolo. La nostra domanda iniziale era «Che cosa si può raccontare ai computer?», dove «raccontare» significa «dare procedure efficaci». E la domanda con cui abbiamo a che fare ora è «È possibile descrivere in termini di procedure efficaci qualunque cosa possa capitarci di voler fare?». Affermare, a questo punto, che ci sono cose che conosciamo ma non sappiamo dire significa non rispondere alla domanda, ma spostare la nostra attenzione dal concetto del *dire*, cui abbiamo cercato finora di ancorarla, a quello del *conoscere*. Come vedremo, questo spostamento è molto opportuno e ha importanza fondamentale; la domanda su cosa possiamo far fare a un computer è, in ultima analisi, una domanda su cosa possiamo far conoscere a un computer. È questa che ci terrà occupati per gran parte del resto di questo libro.

Per il momento, ricordiamo che ho già sollevato quest'argomento; più sopra, ho detto che avere la mappa di una città non significa conoscerla. Analogamente, saper dire le regole degli scacchi non significa sapere giocare a scacchi. Il maestro di scacchi sa più di quanto possa dire. Non sto dicendo (anche se penso sia vero) che non potremo mai trovare il modo di esplicitare tutte le sue conoscenze sugli scacchi; dico soltanto che in questo caso abbiamo un esempio di conoscenza che è efficace pur non essendo per il momento raccontabile. Se fosse vero che la conoscenza degli scacchi di nessun grande maestro è completamente esplicitabile, vorrebbe dire che nessun computer sarà mai in grado di giocare a scacchi al livello dei grandi maestri? Niente affatto. Come ho detto, affronteremo questo genere di problemi nel seguito del libro.

[5] M. Polanyi, *La conoscenza inespressa*, Armando, Roma 1979, p. 4.

3
COME FUNZIONANO I COMPUTER

Da un punto di vista strettamente formale, i computer moderni sono semplicemente delle macchine di Turing che operano su un alfabeto formato dai due simboli «0» e «1», e che possono assumere un numero astronomico di stati. Ma dire questo è come dire che, dal momento che tanto le biciclette quanto gli aeroplani sono veicoli destinati a trasportare persone, essi sono formalmente identici. Il computer moderno differisce dalle macchine di Turing che abbiamo discusso sia per come viene costruito, sia per come viene programmato.

Molte persone sanno che un computer può confrontare il loro nome stampato sulle carte di credito con una lista di nomi immagazzinata, in qualche modo, nel computer stesso. Eppure, la gente per lo più crede che i computer siano fondamentalmente macchine per fare dell'aritmetica su larga scala, cioè nient'altro che una versione ingrandita e molto più veloce dei calcolatori tascabili. Benché quest'opinione sia difendibile sul piano strettamente formale, è molto più utile riconoscere che il computer è fondamentalmente un manipolatore di simboli. Tra i simboli che può manipolare ce ne sono alcuni che gli uomini, e in un certo senso anche i computer, interpretano come numeri. Tuttavia, quasi tutti i computer passano molto, anzi la maggior parte, del loro tempo facendo lavoro non numerico.

Per giustificare quanto ho appena detto, devo dire qualcosa sui simboli e sulla loro interpretazione. E per fare questo, devo anche spiegare come i simboli possano essere rappresentati, specialmente all'interno dei computer.

Almeno per ora, limitiamo la nostra attenzione ai simboli usati per comporre un testo scritto. Ci sono lettere maiuscole e minuscole dell'alfabeto, segni di punteggiatura e simboli speciali come quelli usati in matematica, per esempio parentesi e segni di addizione e di uguaglianza. Lo spazio bianco va considerato anch'esso come un particolare simbolo. Se così non fosse, sarebbe difficile scrivere frasi formate da parole individuali. Libri formati da sequenze di questi soli simboli ci possono far ridere e piangere, possono raccontarci la storia della filosofia, di una persona, di una nazione, e ci possono insegnare molte arti, comprese la matematica. In particolare, ci possono insegnare come costruire algoritmi, cioè procedure efficaci, e ci possono anche dare insiemi di regole che costituiscono algoritmi. Dunque, per quanto informale possa essere la definizione che vogliamo adottare per il concetto di informazione, dob-

biamo riconoscere che i simboli che intendiamo discutere qui sono capaci di trasportare informazione. Come vengono rappresentati e manipolati i simboli nei computer?

Supponiamo che l'alfabeto di cui ci vogliamo interessare consista di 256 simboli distinti; il numero è certamente sufficiente per comprendere tutti i simboli cui ho fatto allusione sopra. Immaginiamo di avere un mazzo di 256 carte, su ciascuna delle quali sia stampato un particolare simbolo del nostro alfabeto, in modo tale che ogni simbolo abbia una carta corrispondente. Se prendiamo a caso una carta dal mazzo, quante domande (cui si possa rispondere «sì» o «no») è necessario fare per sapere qual è il simbolo stampato su quella carta? Sicuramente, il problema si può risolvere facendo al massimo 256 domande. Possiamo decidere un certo ordine per i simboli e iniziare domandando se si tratta del primo simbolo della nostra lista, per esempio chiedendo: «È una A maiuscola?». Se la risposta è «no», chiederemo se si tratta del secondo simbolo, e così via. Ma se ci mettiamo d'accordo con chi ha posto il problema su un certo modo di ordinare i simboli, avremo la possibilità di organizzare le nostre domande in maniera molto più economica. Chiederemo infatti se il simbolo che stiamo cercando è nella prima metà della lista. Qualunque sia la risposta, avremo isolato un insieme di 128 simboli tra i quali c'è quello che cerchiamo. Chiederemo allora di nuovo se esso si trovi nella prima metà dell'insieme ridotto, e così via. Procedendo in questo modo, finiremo per scoprire qual è il simbolo stampato sulla carta scelta dopo esattamente otto domande. Avremmo potuto registrare le risposte ricevute scrivendo «1» quando la risposta era «sì» e «0» quando era «no». Il registro consisterebbe in tal caso di otto cosiddetti *bit* ciascuno dei quali è «1» oppure «0» [1]. Tale sequenza di otto bit è una rappresentazione priva di ambiguità del simbolo che stavamo cercando. Inoltre, una volta stabilito l'ordine dei simboli, ciascuno di essi ha una sola possibile rappresentazione in forma di sequenza di otto bit.

Esistono, in realtà, delle convenzioni per ordinare un insieme di caratteri proprio di questo tipo, e per la loro codificazione individuale. (Il fatto che l'accordo su tali convenzioni, per quanto molto vasto, non sia universale non ci deve preoccupare, almeno per il momento). Qualche anno fa, il particolare schema di codificazione usato nei computer prodotti dall'IBM è diventato qualcosa di molto simile a uno standard per le industrie di tutto il mondo. Secondo questa convenzione una sequenza di otto bit che rappresenta un carattere (di un alfabeto formato da 256 caratteri) è chiamata *byte*, e una catena di quattro *byte* è chiamata *word* (parola).

Abbiamo visto che qualunque testo può essere rappresentato da un riga di 1 e 0. Per poter fare del lavoro utile su informazioni codificate come righe di bit, dobbiamo poterle manipolare secondo qualche criterio preciso, vale a

[1] Se si usa la notazione decimale per i numeri, le *dieci* cifre 0,1,2,....9 vengono in inglese chiamate *digits*. La notazione di cui stiamo discutendo permette però di usare soltanto i *due* simboli «0» e «1», che in inglese vengono perciò chiamati *bits*; in italiano la dizione più corretta sarebbe forse «cifre binarie» (contrapposte a «cifre decimali»). Tuttavia qui, come in altre parti del capitolo, manteniamo la notazione inglese, in quanto essa è ormai praticamente universale (NdT).

dire giocarci secondo certe regole. A questo punto, conosciamo le pedine dei giochi che possiamo fare. Restano da stabilire le regole. Ma prima di arrivarci, lasciatemi dire due parole sulla rappresentazione e manipolazione elettrica di righe di bit.

Un qualunque filo elettrico ha due possibili stati, a seconda che ci passi o meno corrente. Consideriamo un filo collegato a un generatore e a un interruttore opportunamente inserito nel circuito. Quando l'interruttore è chiuso, nel filo passa corrente, altimenti no. Supponiamo che l'interruttore sia a sua volta collegato a un meccanismo che lo apre e lo chiude regolarmente; per esempio, l'interruttore è chiuso per un secondo, poi aperto per un altro secondo, e così via. Possiamo allora parlare del flusso di elettricità nel filo come di un treno di impulsi (vedere Figura 3.1). Un comune campanello da porta d'ingresso è un generatore di impulsi. Quando gli si fornisce potenza, cioè quando il bottone è premuto, un interruttore si chiude. La corrente che passa in un filo fa muovere il martelletto del campanello, aprendo così l'interruttore. Questo viene quindi chiuso di nuovo, e così via.

Un computer moderno, naturalmente, è in sostanza un apparecchio elettrico, proprio come un campanello elettrico. Quando schiacciamo il pulsante di un campanello, abbiamo l'idea che il campanello sia «acceso» anche se, in un certo senso, esso si accende e spegne in continuazione mentre il pulsante rimane schiacciato. In modo simile, si può pensare a un computer in funzione come a un apparecchio acceso, che viene in realtà continuamente acceso e spento da un treno di impulsi come quello che abbiamo discusso. Concettualmente, dunque, il tempo di un computer è diviso in due tipi di intervalli, un intervallo di riposo durante cui esso è, in un certo senso, «spento», e un intervallo attivo durante il quale deve avvenire tutto ciò che si vuole far avvenire. In realtà, il treno di impulsi regolare che abbiamo discusso funge da orologio. Uno stato di «niente corrente» nel filo che trasporta quel treno segnala il periodo di riposo, mentre il passaggio di corrente segnala il periodo attivo.

Per concettualizzare ciò che succede all'interno di un computer, pensate a una mappa ferroviaria che rappresenti l'intera rete di un continente, e in cui le linee ferroviarie siano rappresentate da fili elettrici. Ogni stazione è rappresentata da un paio di lampade al neon. Durante un intervallo di riposo, noteremo che alcune lampade sono accese, altre spente. Durante il periodo attivo successivo, alcune delle lampade accese si spengono, alcune di quelle spente si accendono, altre restano com'erano. Tutto qui. Poi c'è un altro periodo di riposo. Durante il periodo attivo, nessuna lampadina lampeggia: ciascuna di esse rimane com'era, oppure cambia il suo stato esattamente una volta, vale a dire si spegne se era accesa, o si accende se era spenta.

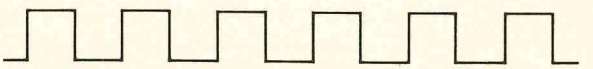


Figura 3.1
Un treno d'impulsi di onde quadrate

L'apparecchio che abbiamo immaginato come formato da una coppia di lampadine al neon è un circuito elettronico che consiste in due componenti identici. Ciascun componente è in grado di far circolare corrente indefinitamente, cioè in pratica di *trattenere* un impulso. Tuttavia, soltanto un componente della coppia può contenere corrente in un dato istante. Due fili entrano nell'apparecchio, uno per ciascun componente. Quando, durante un periodo attivo, un filo trasmette un impulso a un componente in cui, in quell'istante, non circola corrente, una corrente viene indotta in questo componente mentre scompare dall'altra metà dell'apparecchio. Questo apparecchio è dunque capace di saltellare tra due stati; o una delle sue metà è «accesa» e l'altra è «spenta», o viceversa. Viene perciò chiamato *flip-flop* [2].

Ciascuno dei suoi due componenti ha anche un filo in partenza, e ogni filo naturalmente trasporterà corrente (cioè un impulso) durante un periodo attivo quando la corrispondente metà del flip-flop è «accesa». La funzione di un flip-flop in un circuito di computer è quella di «ricordare» su quale delle sue metà è arrivato l'ultimo impulso. Si tratta cioè di un apparecchio per immagazzinare un bit d'informazione.

Adesso abbiamo un altro modo per descrivere cosa succede dentro un computer durante un periodo attivo: molti flip-flop cambiano stato. Ma il compito del computer è manipolare l'informazione, non soltanto trasmetterla da un posto a un altro. E la manipolazione dell'informazione è, come abbiamo già osservato, fondamentalmente una questione di trasformazione. Ora, ogni singolo filo che porta da un lato di un flip-flop a quello di un altro può trasportare al più un impulso durante un intervallo attivo. Perciò, qualunque trasformazione si voglia ottenere durante un intervallo attivo deve essere il risultato di operazioni elettriche non su un flusso di impulsi che si susseguono nel tempo, bensì su una serie di impulsi che avanzano in parallelo.

Immaginate un'enorme rete telefonica in cui ciascun telefono sia collegato in modo permanente con un certo numero di altri telefoni; non ci sono tastiere per fare i numeri. Tutti gli abbonati passano il loro tempo guardando lo stesso canale televisivo, e ogni volta che comincia una pubblicità, cioè un intervallo attivo, prendono la cornetta e gridano «uno» o «zero», a seconda di cosa c'è scritto su un blocco per appunti attaccato al loro apparecchio. Allo stesso tempo, essi ascoltano ciò che viene loro gridato attraverso il telefono, e scrivono «uno» oppure «zero» sul notes, a seconda di ciò che sentono. Un telefono può trasmettere a parecchi ricevitori contemporaneamente, e può anche ricevere da più trasmettitori diversi. Tuttavia, il segnale che raggiunge ciascun ricevitore deve essere, senza ambiguità, un «uno» o uno «zero». Devono perciò esserci dei «centralinisti» (in realtà, sono degli apparecchi elettronici), collocati lungo i fili che collegano i ricevitori fra loro, la cui funzione è quella di comporre un segnale singolo, «uno» oppure «zero», a partire da molti segnali in arrivo. Questi apparecchi sono chiamati «cancelli» (in inglese, *gates*).

Passiamo alla descrizione di tre diversi tipi di cancello, ciascuno con una funzione distinta. Il più semplice è quello che, quando riceve uno «zero», tra-

[2] Espressione inglese, onomatopeica, che significa «saltellare» (NdT).

smette un «uno» e viceversa. Questo è il cancello NOT (la negazione «non», in inglese). La sua funzione è descritta dalle formule

$$NOT(0) = 1$$
$$NOT(1) = 0$$

La sua rappresentazione schematica usuale è mostrata nella Figura 3.2. Ha un ingresso, qui indicato con A, e un'uscita, indicata con B.

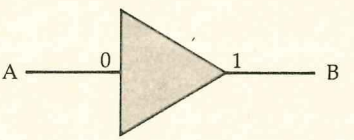


Figura 3.2
Un cancello NOT

Il cancello AND (la congiunzione «e», in inglese) ha due ingressi e una uscita. Trasmette «uno» se e solo se i due ingressi sono entrambi «uno»; altrimenti trasmette «zero». La sua funzione è descritta dalle formule

$$AND\ (0,0) = 0$$
$$AND\ (0,1) = 0$$
$$AND\ (1,0) = 0$$
$$AND\ (1,1) = 1$$

La sua rappresentazione schematica è mostrata in Figura 3.3.

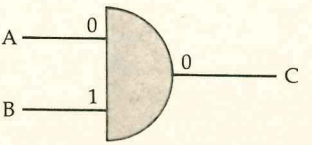


Figura 3.3
Un cancello AND

Il cancello OR (la disgiunzione «oppure», in inglese) ha anch'esso due ingressi e una uscita. Trasmette «uno» ogni qualvolta almeno uno dei suoi ingressi è «uno». Altrimenti, trasmette «zero». Le formule sono:

$$\text{OR } (0,0) = 0$$

$$\text{OR } (0,0) = 1$$

$$\text{OR } (1,0) = 1$$

$$\text{OR } (1,1) = 1$$

La sua rappresentazione schematica è quella della Figura 3.4

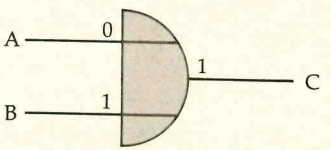


Figura 3.4
Un cancello OR

Il lettore molto motivato potrà divertirsi a seguire il tragitto di impulsi attraverso il sistema di cancelli mostrato nella Figura 3.5. Esso rappresenta un circuito i cui componenti sono cancelli AND, OR e NOT, e la cui funzione è fare la somma aritmetica di due cifre binarie.

(L'addizione binaria è simile a quella decimale, soltanto molto più semplice. La somma decimale di 2 e 3 è 5. La somma decimale di 7 e 8 è anch'essa 5, ma c'è anche un «riporto» di 1 che viene aggiunto nella prima colonna a sinistra. Nell'aritmetica binaria ci sono soltanto due cifre, 0 e 1. Le regole per l'addizione binaria sono molto semplici, per l'esattezza:

$$0+0 = 1$$

$$0+1 = 1$$

$$1+0 = 1$$

$$1+1 = 0 \text{ e riporto di } 1$$

Il «sommatore» binario a un bit mostrato in Figura 3.5 ha perciò tre ingressi A,B, e C e due uscite, S e D. A e B rappresentano i due bit da sommare, C un riporto che può essere stato prodotto da un sommatore simile posto, per così dire, a destra di quello mostrato. S è la somma ottenuta, e D il riporto).

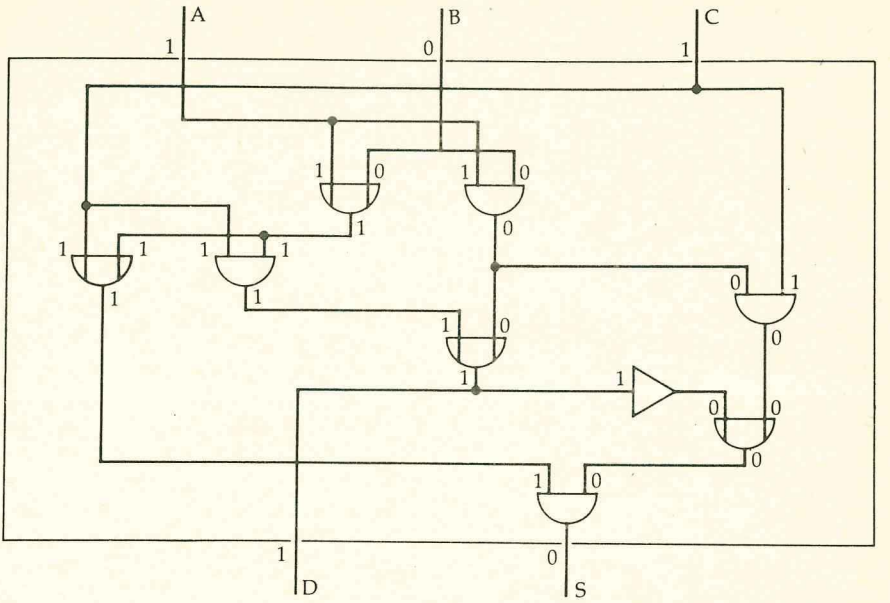


Figura 3.5
Un sommatore binario a un bit con riporto d'ingresso e d'uscita (Figure 3.2 - 3.5 da D.C. Evans, «Computer Logic and Memory», Copyright © 1966 by Scientific American, Inc. All rights reserved.)

Non è in realtà necessario che il lettore segua il percorso di impulsi attraverso questo sommatore. La cosa importante è comprendere che combinazioni dei tre semplici cancelli che abbiamo descritto permettono di mettere in pratica regole di trasformazione sull'informazione. Notate anche che l'intero circuito è stato rinchiuso in una cornice. Questa ha tre ingressi (A,B e C) e due uscite (S e D), ed è essa stessa un'unità. Abbiamo combinato componenti semplici a formare un componente più complesso. Potremmo ora combinare, mettiamo, 32 di questi sommatori e formare un sommatore a 32 bit. Esso sarebbe a quel punto un componente singolo. Tanto la costruzione quanto l'istruzione dei computer non sono altro che procedimenti di questo tipo, in cui componenti piccoli sono riuniti a formare componenti più grandi.

Ritorniamo per un momento alla nostra immagine dei telefoni intercomunicanti, questa volta però tenendo conto del fatto che i cancelli intervengono nelle conversazioni (se possiamo chiamarle così) tra abbonati. Possiamo ora immaginare che i tre abbonati A,B e C alzino la cornetta, al momento giusto, e gridino rispettivamente «1», «0» e «0». Se sono collegati agli abbonati S e D dal circuito che abbiamo mostrato, S sentirà «1» e D sentirà «0», ed entrambi scriveranno il numero sentito sul loro notes. Naturalmente, nei compu-

ter veri tutti questi «abbonati» sono in realtà dei flip-flop, e i periodi di trasmissione sono estremamente brevi. Nei computer, dunque, i risultati dei calcoli — vale a dire delle trasformazioni di informazione — effettuati durante i periodi attivi sono conservati nei flip-flop. Qui essi passano i periodi di riposo, per poter essere disponibili, durante i periodi attivi seguenti, per ulteriori trasformazioni.

Ho già suggerito l'idea che sommatori a un bit possano essere combinati per formare, mettiamo, sommatori a 32 bit. Questi ultimi avrebbero come input due sistemi di 32 flip-flop — in un sommatore a più passaggi, i riporti sono interni — e come output un altro insieme di 32 flip-flop. Si può considerare che ciascuno di questi insiemi contenga, durante i periodi di riposo, un numero binario a 32 bit. Dovrà naturalmente trattarsi di insiemi ordinati, in cui cioè ci sia un primo bit, un secondo bit, e così via. Un simile insieme di flip-flop è chiamato *registro*. È un altro esempio di aggregazione di più componenti elementari. Ma è possibile in generale dire che ogni registro a *n*bit contiene un numero formato da *n*bit? No, o almeno non in modo incondizionato, perché l'interpretazione di una riga di bit contenuta in un registro dipende da quali componenti sono ad esso collegati, e quindi da quali operazioni si possono effettuare sul suo contenuto. Se una particolare coppia di registri è collegata soltanto a componenti che effettuano operazioni aritmetiche, vale a dire se il computer tratta l'informazione ivi contenuta soltanto come una coppia di numeri, allora essi sono effettivamente dei numeri, almeno durante la loro manipolazione all'interno del computer. Per accorgersi che i simboli usati nel linguaggio naturale si comportano in modo simile, basta che uno rilegga con attenzione questa stessa frase, in cui la parola «uno» è di volta in volta un numero, una parola e una sequenza di simboli non interpretata.

Un computer in funzione è impegnato a giocare un gioco elaborato e molto complicato. Dopo ogni intervallo di riposo, fa una mossa. Il contenuto di molti registri è trasportato in altri, più o meno come i pezzi degli scacchi vengono spostati sulla scacchiera. Nel computer, però, moltissimi pezzi (bit) vengono mossi contemporaneamente e, ciò che è più importante, bit individuali e insiemi di bit vengono trasformati durante il viaggio da un posto all'altro.

Ho già ricordato il fatto che, in molti computer, insiemi di otto bit sono aggregati a formare dei byte. Un registro a otto bit è una realizzazione fisica di questa aggregazione concettuale. I computer contengono molti registri collegati tra loro da reti di cancelli del tipo che ho descritto. L'intero insieme di questi registri, e della rete logica che li unisce, costituisce in un certo senso il computer. Questo insieme è la macchina, spesso chiamata unità centrale di elaborazione del computer (o CPU, dall'inglese *central processing unit*), che effettua la vera e propria manipolazione di simboli logici durante ciascuno dei periodi attivi del computer. La configurazione totale degli stati dei suoi singoli flip-flop costituisce lo stato del computer, in un senso molto simile a quello che intendevo quando, nei capitoli precedenti, ho parlato dello stato di una macchina di Turing.

Una macchina di Turing del tipo discusso in precedenza riceve da un nastro magnetico l'informazione su cui deve operare. Deve leggere da e scrivere

su questo nastro in modo sequenziale, un simbolo per volta. Un computer moderno, invece, conserva al suo interno gran parte dell'informazione che manipola. Il sistema di immagazzinamento interno del computer consiste in un enorme sistema di registri a otto bit, un milione o anche più. Questi sono disposti in un ordine preciso, un po' come le cassette delle lettere nell'ingresso di un grande condominio. I registri sono numerati in modo sequenziale, a cominciare da 0,1, e così via. Il numero di ciascun registro è chiamato il suo *indirizzo*. Poiché ciascun registro ha un unico indirizzo, si può parlare dell'indirizzo di un registro in modo del tutto indipendente dal suo contenuto, cioè dallo stato dei suoi otto flip-flop, e viceversa. Immaginate ora una griglia di cassette delle lettere in un condominio, che sia dotata di un lucchetto a combinazione. Per poter prendere qualcosa da una particolare cassetta, o per poterci mettere qualcosa, bisogna prima specificarne l'indirizzo nel lucchetto. Il magazzino del computer ha per l'appunto un apparecchio simile. È, naturalmente, un registro (un registro a 20 bit è sufficiente per dare un indirizzo a 1.048.576 cassette). È comodo pensare a questo registro degli indirizzi come a una parte della CPU del computer, benché esso sia collegato tanto alla CPU quanto alla memoria del computer. Ma vederlo come parte della CPU aiuta a visualizzare il fatto che il suo contenuto è esso stesso manipolabile dalla macchina. Esso può, per esempio, contenere risultati intermedi in qualche lunga sequenza di calcoli aritmetici che possono essere usati, in pratica, come numeri ordinali.

L'indirizzabilità della memoria di un computer moderno è una delle sue proprietà più importanti. Per apprezzare l'enorme differenza che essa fa nel modo di fare ricerche in memoria, considerate il seguente problema. Una città ha meno di 10.000 abbonati del telefono. La guida del telefono di questa città elenca tutti gli abbonati in ordine alfabetico e, naturalmente, ne dà i numeri di telefono. Ma essa elenca anche il numero di serie di ciascun abbonato nella guida stessa. Per esempio:

1 Abatantuono, Achille 67.44.11

.
.
.

1423 Invernizzi, Elvira 58.53.24

Ora, vogliamo sapere se esistono coppie di utenti tali che le ultime quattro cifre del numero di telefono dell'uno corrispondono al numero di serie dell'altro, e viceversa. Dato l'elenco esemplificato sopra, esiste per esempio un utente

5324 Strillone, Vincenzo 75.14.23

corrispondente in quel senso a Elvira Invernizzi?

Un modo semplice di risolvere il problema consiste nello scorrere l'elenco, a cominciare dal primo utente, e nel controllare se l'utente il cui numero di serie è dato dalle ultime quattro cifre del numero di telefono soddisfa alla condizione posta. La risposta al problema è «sì» se si riesce a trovare una coppia di questo tipo, mentre è «no» se, dopo aver controllato tutto l'elenco, non se ne sono trovate. Il caso peggiore che si può incontrare è quello in cui coppie simili non esistono. In tal caso, ciascun utente verrà controllato al massimo due volte. Ma supponiamo che la guida telefonica fosse registrata su un nastro del tipo richiesto dalle macchine di Turing che abbiamo descritto. Allora, anche trascurando il movimento del nastro associato con i compiti di «contabilità», sarebbe necessario incontrare ciascun utente molto più di due volte. Nel nostro esempio, già soltanto il primo controllo richiederebbe di guardare il primo utente, i 4409 utenti tra questo e il 4411esimo, e il 4411esimo stesso. Nelle macchine di Turing classiche, il nastro non verrebbe semplicemente fatto passare sulle voci intermedie, ma ciascuna dovrebbe proprio essere controllata e interpretata. È lo stesso principio che fa arrabbiare quanti sanno di essere citati in un grande volume che, purtroppo, è sprovvisto di indice. Si trovano di fronte alla prospettiva di doverlo leggere tutto.

Questo esempio della guida telefonica mostra non tanto che l'indirizzabilità aumenta l'efficienza di molte ricerche, quanto che essa, in certe circostanze, le rende proprio non necessarie. Infatti, se avessimo scelto un particolare utente e ci fossimo chiesti se era possibile accoppiarlo a un altro nel modo indicato, sarebbe stato possibile rispondere direttamente, senza scrutinare alcun dato inessenziale. Per esempio, il signor Abatantuono fa o meno parte di una coppia? Per scoprirlo, andiamo direttamente a controllare l'utente numero 4411. Se le ultime quattro cifre del suo numero di telefono sono «0001», allora sì; altrimenti no.

Tenendo presente tutto questo, ritorniamo alle quintuple che definiscono una macchina di Turing *T*, costituendo così un programma per una macchina di Turing universale che ha il compito di imitare *T*. Come ricorderete, la forma generale di una quintupla di questo tipo è:

(stato attuale,
simbolo attuale,
prossimo stato,
nuovo simbolo,
direzione di moto).

In generale, una macchina di Turing che si trova in un certo stato, diciamo 19, e scrutina un certo simbolo, diciamo «1», deve passare in rassegna tutte le quintuple del suo programma sempre chiedendosi, per così dire, se la particolare quintupla (regola) che sta leggendo è quella che si applica allo stato 19. Quando trova quella corrispondente al suo stato e al simbolo che stava scrutinando, riscrive quel simbolo, sposta il nastro e (se è il caso) cambia stato. A questo punto, la ricerca ricomincia. Se, tuttavia, le quintuple fossero registrate in una memoria a indirizzi, la ricerca della quintupla appropriata po-

trebbe essere evitata, o per lo meno accorciata. Supponiamo, per esempio, che tutte le quintuple associate con un particolare stato occupino 100 byte di spazio di memoria, e che la prima sia memorizzata a cominciare dal registro 1000. Allora l'insieme corrispondente allo stato 19 sarebbe memorizzato a cominciare dal registro numero 1000 + (19-1)x100, cioè dal registro 2800. La nozione di stato viene così trasformata in quella di indirizzo, almeno in questo contesto.

Nei computer reali, anche i dati sono registrati nella memoria a indirizzi del computer. Questa innovazione ci permette di eliminare anche la restrizione per cui soltanto un dato immediatamente vicino a quello attualmente scrutinato può essere immediatamente accessibile. Per cominciare a vedere come l'indirizzabilità venga usata nella composizione dei programmi per computer, esaminiamo un problema piccolo ma realistico.

Vogliamo calcolare radici quadrate di numeri (vale a dire, dato un numero, per esempio 25, vogliamo sapere qual è il numero che, moltiplicato per se stesso, dà quel numero. La radice quadrata di 25 è 5, poichè 5x5 = 25). Facciamo finta di avere un fedele servitore (umano) che instancabilmente obbedirà a qualunque nostra istruzione. Conosciamo un algoritmo per calcolare la radice quadrata di qualunque numero positivo. Dato il numero *n*, ne calcoliamo la radice quadrata come segue:

1. Per prima cosa facciamo un'ipotesi, sempre la stessa, e cioè che la radice sia 1.
2. Poi arriviamo a un'ipotesi migliore, come segue:
   (a) moltiplichiamo la vecchia ipotesi per se stessa;
   (b) aggiungiamo al risultato il dato numero *n*;
   (c) dividiamo tale somma per il doppio della vecchia ipotesi
   (che questo calcolo porta sempre a una approssimazione migliore è stato provato niente meno che da Isaac Newton; a meno che, naturalmente, la vecchia ipotesi non fosse già la risposta giusta).
3. Se la differenza tra la vecchia ipotesi e quella migliore ci sembra abbastanza piccola, allora accettiamo la nuova approssimazione. Altrimenti, ne calcoleremo una ancora migliore.

Il procedimento si giustifica nel modo seguente.

Supponiamo di voler trovare la radice quadrata di 25, e che la nostra ipotesi iniziale sia, mettiamo, 4. Sappiamo che questo numero è troppo piccolo, poichè 4x4 = 16. Se dividiamo il numero dato inizialmente, cioè 25, per la nostra ipotesi, allora il risultato sarà più grande del risultato che stiamo cercando, 5 in questo esempio. Questo quoziente, e l'ipotesi che abbiamo fatto, approssimano perciò da parti opposte il risultato che cerchiamo. Se dunque facciamo una media di questi due, otterremo un'altra approssimazione, e per di più migliore di quella da cui siamo partiti. Possiamo poi iterare questa formula finché otterremo un risultato vicino quanto vogliamo a quello esatto.

Il nostro servitore, però, non è troppo intelligente. Ha lavorato soltanto con moduli per le tasse che richiedono di calcolare questo e quest'altro, e di scriverlo nella tal riga o nel tal riquadro. Perciò, abbiamo compilato il formu-

lario mostrato nella Tabella 3.1 [a p. 94, NdR], basato proprio su quel tipo di procedure. Il nostro servitore lavora in una stanzetta con una feritoia per i messaggi in entrata e una per quelli in uscita. Appena un foglietto con un numero scritto sopra compare nella feritoia di ingresso, egli inizia a lavorare furiosamente. Quando ha finito, scrive il risultato su un altro foglietto che infila nella feritoia in uscita. La sua unica istruzione iniziale è quella di cominciare obbedendo all'istruzione 101 del formulario e, a meno di trovare un ordine contrario nel formulario stesso, di continuare a seguire le istruzioni nell'ordine in cui sono scritte. Tra l'altro, farebbe bene a scrivere soltanto a matita, e ad avere una buona gomma a portata di mano; ha infatti poco spazio a disposizione sul formulario, per scrivere più e più volte. (Non è il caso che il lettore esegua un calcolo esemplificativo basato sul formulario proprio fino in fondo. Può essere utile, tuttavia, procedere abbastanza da generare, diciamo, due migliori approssimazioni per la radice quadrata di 25).

Notate l'importante ruolo svolto dalla casella A. Essa è in realtà un registro che contiene risultati intermedi di calcoli. Notate anche che nessuna istruzione individuale si riferisce a più di un indirizzo di riga. Istruzioni dotate di questa proprietà sono chiamate *a indirizzo singolo*. Possiamo effettuare, per esempio, delle addizioni (che naturalmente richiedono due operandi) cominciando col memorizzare un operando nella casella A, poi aggiungendo l'altro al contenuto della casella A e lasciando di nuovo la somma nella casella A. Le righe 121, 123, e 124 servono come registri di memoria temporanea. Infine, osservate che le istruzioni e la memorizzazione dei risultati intermedi sono organizzati in modo tale che un formulario come quello della Tabella 3.1 può essere usato più volte. A ogni nuovo uso, la casella A e le righe 121, 123, e 124 conterranno numeri irrilevanti per il nuovo obiettivo da raggiungere. Ma quei numeri non interferiranno con il nuovo calcolo. Infatti, abbiamo usato una consuetudine universalmente accettata da chi scrive programmi: laddove in molti bagni pubblici è esposto un cartello che invita gli utenti a lasciare tutto nelle condizioni in cui l'avevano trovato, noi adottiamo proprio la convenzione opposta. Diciamo: «Lasciate tutto nelle condizioni in cui volete che sia al momento di proseguire il lavoro. Per esempio, memorizziamo sempre la «vecchia approssimazione» nella riga 123. Dopo l'uso del formulario, quella riga conterrà l'ultima «vecchia approssimazione» dell'ultimo calcolo portato a termine. Ma un calcolo nuovo, per cui l'ultima «vecchia approssimazione» memorizzata è del tutto irrilevante, comincia copiando «1», cioè il primo tentativo standard, nella riga 123 (vedere le righe 102 e 103).

Per qualche ragione che — a quanto pare — nessuno sa più, un procedimento di calcolo come quello che stiamo discutendo viene chiamato *routine*. Quando si inizia un calcolo come questo, si dice che si «entra nella routine», e quando si è completato l'ultimo passo si «esce dalla routine». Ma routine piccole come quella illustrata qui non sono quasi mai di grande utilità in se stesse; quante volte ci capita di voler sapere la radice quadrata di un numero, al di fuori del contesto stabilito da un qualche calcolo più impegnativo? Questo calcolo più impegnativo può essere, mettiamo, il calcolo delle radici di un'equazione quadratica, cioè un'equazione che ha la forma generale

$$ax^2 + bx + c = 0$$

Le radici cercate sono date dalla formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

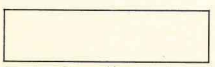
Questo dunque è un problema — trovare le radici di una data equazione quadratica — che implica come sottoproblema il calcolo della radice quadrata di un numero. È possibile scrivere un formulario molto simile alla Tabella 3.1 per istruire un computer umano nella soluzione del problema complessivo. Però, dal momento che già abbiamo un formulario che indica come calcolare le radici quadrate, faremo meglio a dire alla persona che lavora sul problema più complesso di ricorrere al formulario già preparato per risolvere il sottoproblema, quando ci arriverà, vale a dire una volta calcolata la quantità $(b^2 - 4ac)$. La routine per calcolare le radici quadrate diventa allora una *subroutine* della routine più complessa. Naturalmente, quest'ultima potrà a sua volta essere una subroutine di una routine ancora più complessa, e così via. Di nuovo, voglio sottolineare che i computer e il lavoro di calcolo hanno per lo più a che fare con la costruzione di strutture gerarchiche più grandi a partire da strutture più piccole. Ed è proprio perché la routine delle radici quadrate che abbiamo mostrato può essere usata come subroutine in una procedura più grande che abbiamo fatto in modo che si inizializzasse da sola (vale a dire, che si «rimettesse in ordine») prima di ogni uso.

Il lettore più attento avrà notato che ho trascurato un problema di gestione che certamente ci verrà a tormentare, se i nostri servitori da calcolo sono davvero così sempliciotti, obbedienti alla lettera e privi di fantasia come li ho dipinti, nell'interpretare le nostre istruzioni. Abbiamo supposto, e continueremo a supporre, che un lavoratore sappia come sommare un numero letto su una riga al contenuto della casella A, come lasciare la somma nella casella A, come riconoscere se un numero registrato nella casella A sia maggiore di zero, e così via. Abbiamo supposto, in altre parole, che egli sappia quanto serve a interpretare correttamente le istruzioni del formulario per le radici quadrate. Ma come facciamo a insegnare a una persona di simili capacità, che però lavora su un problema più complesso, a subappaltare (per così dire) i suoi sottoproblemi?

È un problema che possiamo evitare di affrontare, semplicemente supponendo che il nostro lavorante sappia come calcolare, diciamo, le radici quadrate nello stesso modo in cui supponiamo che sappia sommare e sottrarre. Ma dovremo anche supporre che sappia già come calcolare le radici di un'equazione quadratica? Se non smettiamo di fare ipotesi di questo genere, finiremo per ammettere di non avere affatto bisogno dei computer, perché saremo giunti ad affermare che, per scrivere un formulario corrispondente a un qualsiasi calcolo, tutto quello che dobbiamo fare è, in realtà, scrivere «fallo».

Tabella 3.1 *Istruzioni per calcolare radici quadrate*

101	Prendi il numero della feritoia d'ingresso, e scrivilo alla riga 121.
102	Copia il contenuto della riga 122 nella casella A.
103	Copia il contenuto della casella A sulla riga 123.
104	Copia il contenuto della riga 123 nella casella A.
105	Somma il contenuto della riga 123 a quello della casella A, e scrivi il risultato nella casella A.
106	Copia il contenuto della casella A sulla riga 124.
107	Copia il contenuto della riga 123 nella casella A.
108	Moltiplica il contenuto della casella A per il contenuto della riga 123, e scrivi il risultato nella casella A.
109	Somma il contenuto della casella A a quello della riga 121, e scrivi il risultato nella casella A.
110	Dividi il contenuto della casella A per quello della riga 124, e scrivi il risultato nella casella A.
111	Copia il contenuto della casella A nella riga 124.
112	Sottrai il contenuto della riga 123 da quello della casella A, e scrivi il risultato nella casella A.
113	Copia il valore assoluto della casella A nella casella A.
114	Sottrai il contenuto della riga 125 da quello della casella A, e scrivi il risultato nella casella A.
115	Se il contenuto della casella A è maggiore di zero, passa alla riga 118.
116	Metti il contenuto della riga 124 nella feritoia d'uscita.
117	Fermati.
118	Copia il contenuto della riga 124 nella casella A.
119	Copia il contenuto della casella A nella riga 123.
120	Passa alla riga 105.
121	0
122	1.0
123	0
124	0
125	.001



Casella A

Abbiamo magari il diritto di credere che la gente sappia sommare, sottrarre ecc. Se, però, vogliamo che vengano eseguite delle manipolazioni di simboli più complicate, dovremo descrivere le procedure opportune in termini più semplici, già capiti. D'altra parte, non vogliamo essere costretti a descrivere ogni volta una procedura di uso frequente. Ogni routine è potenzialmente una subroutine di qualche routine più grande. Perciò, dopo aver scritto una particolare routine una volta, vogliamo che entri a far parte di una biblioteca di subroutine da cui essa può venir richiamata quando occorre.

Ogni biblioteca ha delle convenzioni per richiedere volumi dagli scaffali; il lettore riempie un modulo di richiesta, lo dà al bibliotecario ecc. La nostra situazione, però, è un po' più complicata di quella di una normale biblioteca per il prestito di libri. Noi vogliamo che il nostro lavorante, oltre a esaminare il formulario corrispondente al compito da eseguire, faccia anche il lavoro. Per esempio, supponiamo che un cuoco abbia preparato una ricetta fino a un certo punto, e qui desideri l'aiuto di uno specialista, che svolga la funzione per cui è famoso sugli ingredienti che il cuoco ha già preparato. Il cuoco può riposarsi mentre lo specialista lavora, e continuerà la sua fatica soltanto dopo che lo specialista gli avrà restituito l'impasto desiderato. Supponendo che ci siano molti di questi specialisti, come fa il cuoco a lasciare ogni volta il controllo della cucina proprio allo specialista giusto, come fa ad accertarsi che lo specialista lavori sugli ingredienti giusti, e come fa infine a riacquistare il controllo della sua cucina, e a sapere dove si trova il risultato della fatica dello specialista? Un modo per fare tutto questo è il seguente.

> Il cuoco lascia tutti gli ingredienti su cui lo specialista deve lavorare in armadietti numerati in modo sequenziale, per esempio il primo nel 119, quello dopo nel 120, e così via. Nell'armadietto che segue quello contenente l'ultimo ingrediente, scrive il numero della stanza in cui starà mentre lo specialista lavora. Lascerà sulla cucina a gas il numero dell'armadietto in cui si trova il primo ingrediente. Infine, suonerà il campanello della stanza dello specialista, di cui naturalmente conoscerà il numero.
> Lo specialista conosce molto bene le convenzioni del cuoco. Quando sente il campanello suonare, va in cucina, trova sulla cucina a gas il numero dell'armadietto con il primo ingrediente, e comincia a fare i suoi trucchi. Quando ha finito, lascia il suo impasto sul gas, suona il campanello della stanza di cui ha trovato il numero nel primo armadietto dopo quelli che contenevano ingredienti, e se ne va.

Perché questo schema funzioni, è necessario solo che sia il cuoco sia tutti gli specialisti che possono venir chiamati conoscano le convenzioni, che il cuoco sappia esattamente di quanti ingredienti ha bisogno lo specialista che sta per chiamare (perché non succeda che quest'ultimo, che lavora in modo così letterale, metta a bollire in pentola l'appunto con il numero della stanza del cuoco insieme agli altri ingredienti!), e che il cuoco sappia l'indirizzo del particolare specialista che vuole chiamare.

Possiamo adottare convenzioni sostanzialmente identiche per chiamare le subroutine nella nostra struttura a formulario. Considerate, per esempio, il se-

guente frammento di formulario, che è parte di una routine più generale. Nel momento in cui il frammento comincia, un numero di cui si vuole calcolare la radice quadrata si trova nella casella A. Questo frammento chiama la subroutine delle radici quadrate, e accerta anche che il controllo del calcolo ritorni alla routine principale.

508. Copia il contenuto della riga 680 nella casella A.
(*Nota:* la riga 680 contiene il numero di cui si vuole calcolare la radice quadrata).
509. Copia il contenuto della casella A nella riga 512.
(*Nota:* la riga 512 è l'«armadietto» in cui il primo — in questo esempio, l'unico — «ingrediente» viene conservato).
510. Copia il numero «512» nella casella A.
(*Nota:* la casella A ha il ruolo di memoria. In questo modo, si è detto allo specialista dove cominciare a cercare gli ingredienti).
511. Passa alla riga 98. (*Nota:* la subroutine delle radici quadrate comincia alla riga 98).
512. (*Nota:* questa riga serve come memoria per il numero di cui si vuole la radice quadrata).
513. 514 (*Nota:* questa riga contiene l'indirizzo della riga cui lo specialista deve restituire il controllo una volta finito il suo lavoro).
514. (sarà l'istruzione seguente nella routine più generale)
(*Nota:* quando si arriva a questa istruzione, la radice quadrata del numero memorizzato alla riga 680 si trova nella casella A).

Il formulario dello specialista è già stato mostrato nella Tabella 3.1. Tuttavia, dobbiamo modificarlo per tener conto delle convenzioni che stiamo discutendo. La versione originaria dice al lavorante di prendere il numero su cui deve operare da una «feritoia d'ingresso». Ora, naturalmente, il numero della riga che contiene quel dato si trova nella casella A. Perciò scriveremo:

98. Copia il contenuto della casella A nella riga 126.
99. Copia il contenuto della riga il cui numero è contenuto nella riga 126, nella riga 121.
(*Nota:* questo è ciò che si chiama «indirizzare in modo indiretto». La routine era originariamente organizzata in modo da operare su un numero memorizzato nella riga 121. In questo modo, possiamo lasciare immutata la routine).
100. Aggiungi «1» al contenuto della casella A, e lascia la somma nella casella A stessa.
(*Nota:* il contenuto della casella A è immutato dall'esecuzione delle istruzioni 98 e 99. Quindi la casella A contiene ora «513», cioè il numero di riga che, nella routine che chiama, contiene l'indirizzo di ritorno; in questo esempio, la riga 513 contiene «514»).
101. Copia il contenuto della casella A nella sezione-indirizzo della riga 117.
(*Nota*: la riga 117 era un'istruzione di fermata, ma noi la sostituiremo come mostrato più avanti. Poiché tutte le istruzioni sono istruzioni a indirizzo singolo, il termine «sezione-indirizzo» è privo di ambiguità quando riferito a istruzioni).

Ora riscriviamo le righe 116 e 117 come segue:

116. Copia il contenuto della riga 124 nella casella A.
(*Nota*: stiamo lasciando l'«impasto», cioè la radice quadrata del numero dato, nella casella A).
117. Passa all'istruzione 0.
(*Nota*: l'indirizzo «0» in questa riga sarà naturalmente sostituito ogniqualvolta si obbedirà all'istruzione 101, in altri termini ogni volta che la subroutine viene chiamata).

Il mio obiettivo era quello di illustrare uno dei molti possibili modi di chiamare una subroutine. Benché questo particolare sistema funzioni, non è efficace come alcuni altri, e non va considerato come standard. Ciò che importa è il principio secondo cui, una volta che una routine è stata scritta, può essere convertita in una componente di una routine più grande. Quest'ultima può a sua volta essere trattata come subroutine di una routine ancora più grande, e così via. Una ragione per cui questo tipo di costruzione gerarchica delle strutture di calcolo (cioè di manipolazione di simboli) è importante, è che esso libera il progettista del sistema di alto livello dall'incombenza di sapere con precisione come (cioè in base a quale algoritmo) funziona la subroutine di livello inferiore. Egli deve sapere come raggiungerla, quali sono le convenzioni per la chiamata e qual è il rapporto funzionale tra input e output. Una subroutine è perciò piuttosto simile a una pratica legale talvolta complessa, mettiamo una licenza di costruzione: si riempiono gli spazi vuoti, si firma e si archivia. Di solito, ciascun firmatario giustamente crede di essersi impegnato a fare certe cose e non altre, e di sapere con esattezza le possibili conseguenze legali. Ma talvolta si finisce nei guai quando ci si affida a moduli già stampati, di cui si crede di capire gli «intenti» ma di cui si trascura di leggere le parti scritte in piccolo. Su questo avrò qualcos'altro da dire più avanti.

Un lavoratore umano il cui compito sia comporre formulari come quelli che ho descritto si stancherebbe presto di dover scrivere così tanto. Inventerebbe delle abbreviazioni. Nessuna informazione essenziale va perduta, per esempio, se la riga

(103. Copia il contenuto della riga 123 nella casella A)

del formulario originale (Tabella 3.1) viene abbreviata con

(103. GET 123)

o se la riga

(105. Aggiungi il contenuto della riga 123 al contenuto della casella A, e scrivi il risultato nella casella A)

viene abbreviata con

(105. ADD 123).

Possiamo inventare abbreviazioni simili (usando, in realtà, verbi della lingua inglese) per ciascuna delle operazioni citate nel formulario originale, e possiamo usarle per codificare l'intera procedura, come mostrato nella Tabella 3.2.

Tabella 3.2 *Una codificazione della routine per radici quadrate*

RIGA	COMANDO		COMMENTO
101	INP	121	Prendi il numero d'ingresso e mettilo in 121.
102	GET	122	Prendi il tentativo iniziale.
103	STO	123	Memorizza come «vecchio tentativo» in 123.
104	GET	123	Prendi il «vecchio tentativo».
→105	ADD	123	Raddoppialo.
106	STO	124	Memorizza «due volte il vecchio tentativo» in 124.
107	GET	123	Prendi il «vecchio tentativo».
108	MPY	123	Moltiplicalo per se stesso.
109	ADD	121	Aggiungi il numero dato.
110	DIV	124	Dividi il risultato per «due volte il vecchio tentativo».
111	STO	124	Memorizza il «nuovo tentativo» in 124.
112	SUB	123	Sottrai il «vecchio tentativo» dal «nuovo tentativo».
113	ABS		Rendi positivo il risultato.
114	SUB	125	Sottrai la tolleranza, data in riga 125.
115	JGZ	118	Se il risultato è positivo, salta i prossimi due passaggi.
116	OUT	124	Produci il «nuovo tentativo» come risultato in uscita.
117	STP		Stop.
→118	GET	124	Prendi il «nuovo tentativo».
119	STO	123	Mettilo al posto del «vecchio tentativo».
120	JMP	105	Ricomincia il «ciclo».
121	0		Posto riservato al numero su cui lavorare.
122	1.0		Posto riservato al «vecchio tentativo» iniziale.
123	0		Posto riservato al «vecchio tentativo».
124	0		Posto riservato al «nuovo tentativo».
125	.001		Posto riservato alla tolleranza.

La procedura, così codificata, è ancora leggibile da una persona. Ciò che più ci importa in questo momento, tuttavia, è che un simile codice è soprattutto leggibile da un computer. Non che il formulario originale, scritto in italiano, non lo sia. Dopo tutto, abbiamo visto che qualunque testo può essere ridotto a sequenze di 0 e 1. Ma il codice mostrato nella Tabella 3.2, privo dei suoi commenti, ha un formato rigido, ed è perciò facile da decodificare. Ogni riga è un comando dato al computer, e ogni comando di questo tipo consiste di due componenti: un'operazione da eseguire, per esempio ADD, e l'indirizzo di un operando. La casella A non viene più citata esplicitamente. È sottinteso che la casella A contiene sempre l'operando implicito o non citato, quando ne serve uno, e che il risultato di qualsiasi operazione, per esempio la somma ottenuta facendo un'addizione, è memorizzato lì. Ma dal punto di vista del computer, a differenza che da quello umano, questo codice è ancora di gran lunga troppo prolisso. La sezione di un comando che codifica l'operazione, ad esempio ADD, può essere sostituita da un carattere che occupi un solo bit (ricordate che ci sono 256 caratteri disponibili). L'unica restrizione imposta allora da una simile convenzione è che non possiamo richiedere che esistano più di 256 operazioni cosidette «interne» (*built-in* in inglese, NdT), vale a dire operazioni che non richiedono ulteriori spiegazioni al lavoratore. Tali operazioni vengono chiamate primitive.

Per continuare la conversione di questa procedura esemplificativa in codice della macchina, supponiamo che il codice, della lunghezza di un byte, per ADD sia «00110101». Allora, se la lunghezza di un comando memorizzato in un particolare computer è di quattro byte, cioè 32 bit, il codice del computer per la riga 105 del nostro formulario sarebbe

00110101000000000000000000001111011

L'intero codice per la nostra routine sarebbe una serie di 25 righe di questo tipo.

Abbiamo illustrato come sia possibile organizzare un lavoro di calcolo in forma di procedura eseguibile da esseri umani, e come tale procedura possa essere tradotta in una notazione che, benché criptica, può facilmente essere gestita da un computer. Ma abbiamo truffato, almeno un pochino. Il nostro consiglio al fedele servitore-calcolatore era di obbedire alle istruzioni del formulario nell'ordine in cui sono scritte, a meno che una particolare istruzione non gli dica di fare altrimenti. Abbiamo dovuto supporre, naturalmente, che egli comprendesse e sapesse come eseguire le operazioni primitive citate nel formulario. Ma abbiamo anche tacitamente supposto che sapesse ricordare la sua posizione sul formulario anche se distratto, mettiamo, dall'esecuzione di una lunga divisione. In realtà, seguire il flusso del controllo in un programma è un compito ben distinto e separabile da quello di eseguire operazioni primitive. Dovremmo dunque aggiungere un'altra casella al formulario, una casella P *(contatore del programma)* che in ogni istante contenga il numero di riga (cioè l'indirizzo) dell'istruzione che viene eseguita in quel momento. Dovrem-

mo poi nominare un supervisore, il cui compito è dire al lavoratore instancabile su quale riga troverà la sua istruzione, vale a dire qual è la riga il cui indirizzo è memorizzato nella casella P. Quando il lavoratore dice al supervisore che l'istruzione assegnata è stata eseguita, il supervisore aggiunge «1» al contenuto della casella P, memorizza nella stessa il risultato, e dà al lavoratore la nuova istruzione. L'effetto di un'istruzione che turba il normale flusso sequenziale del controllo è, naturalmente, quello di sostituire il contenuto della casella P con l'indirizzo della nuova istruzione da eseguire. Per esempio, l'effetto dell'esecuzione della riga

(120. Passa alla riga 105)

è che il lavoratore copia il numero «105» nella casella P e immediatamente va a eseguire l'istruzione scritta nella riga 105, senza riferire al superiore di aver finito quell'istruzione.

Nei veri computer, il lavoro è ancor più suddiviso. Ci sono componenti destinati a eseguire ciascuna delle operazioni interne, o primitive, di cui il computer è capace.Questi componenti sono essenzialmente separati, anche se possono in certi casi avere dei circuiti in comune. E ci sono altri componenti che gestiscono l'accesso alla memoria del computer. Molti computer hanno addirittura dei piccoli sotto-computer specializzati, la cui unica funzione è quella di controllare il trasferimento di informazione dal mondo esterno al computer principale. Se discutessi qui tutti questi dettagli, mi allontanerei troppo dal mio argomento principale. Ho, tuttavia, evidenziato l'idea di flusso del controllo di un programma, perché è davvero un concetto importante. Notate che a lato del programma mostrato nella Tabella 3.2 sono state tracciate due linee. Esse illustrano il flusso del controllo in quel programma o, per dirla in altri termini, la struttura del programma. Ciò che rende interessante un programma pure così piccolo è il fatto che esso comprende un *ramo condizionale*. Il fatto che il programma, durante la sua esecuzione, passi o meno per un'altra iterazione è determinato dal risultato dell'istruzione della riga 115. Se, quando si arriva a quell'istruzione, la casella A contiene un numero maggiore di zero, si procederà a un'altra iterazione. Altrimenti, il programma si ferma dopo pochi altri passi.

La capacità dei computer di eseguire istruzioni a diramazione condizionale (vale a dire, di modificare il flusso di controllo dei programmi in funzione del risultato di test sui risultati intermedi dei loro stessi calcoli) è una delle loro proprietà più cruciali, perché ciascuna procedura efficace può essere ridotta a una serie di comandi (vale a dire, affermazioni del tipo «fai questo» e «fai quello») collegati a istruzioni a diramazione condizionale. Inoltre, sono necessarie soltanto istruzioni a diramazione binaria (vale a dire, istruzioni del tipo «se questo e quello sono veri, fai questo; altrimenti, fai quello»). Se la decisione se questo e quello siano veri comporta essa stessa procedure complesse, anche queste possono essere messe in forma di comandi e di istruzioni a diramazione binaria (cioè, a due vie).

Ho tracciato finora un quadro piuttosto accurato — benché sia solo un quadro — di cosa un programma è e di come un computer esegue un programma. Nel corso della discussione, abbiamo ridotto un formulario molto prolisso a un codice assai compatto, e quindi abbiamo ridotto quel codice ancora di più, fino ad avere righe di bit strutturate secondo un formato rigoroso. Abbiamo supposto, durante tutto questo processo, che nessuna di queste riduzioni fosse accompagnata dalla perdita di informazioni essenziali. Speriamo, in altre parole, che la riga di 0 e 1 che è il prodotto finale di tutte le nostre trasformazioni abbia lo stesso significato del nostro formulario originale. Per controllare se questo sia vero o meno, dobbiamo sapere cosa ciascuna delle due cose significhi separatamente.

Che cosa significa il formulario che abbiamo costruito? Per un analfabeta, può significare la possibilità di costruire un aeroplano di carta; e per una madre che crede che sia stato il suo bambino di dieci anni a scriverlo, può significare che suo figlio è un genio in embrione. Ma nel nostro contesto lo vediamo come un insieme di istruzioni. Quindi, per noi il suo significato è l'azione intrapresa da qualcuno che capisce quelle istruzioni, se si mette a eseguirle. Il significato di un codice per computer deve dunque essere l'azione che un computer intraprende quando interpreta quel codice come un insieme di istruzioni, e le esegue. Contiamo di sapere che cosa farà una persona che segua le istruzioni scritte sul formulario, perché sono scritte in un linguaggio che entrambi conosciamo. Ciò che farà il computer, dato il codice che abbiamo preparato, dipende interamente da come è stato progettato. Il progetto determina il significato di qualsiasi sequenza di 0 e 1 data a un computer. Se la particolare sequenza di bit che abbiamo elaborato deve avere per un computer lo stesso significato che il nostro formulario ha per un calcolatore umano, allora il computer deve avere componenti equivalenti, dal punto di vista funzionale, a ciascuna delle operazioni che chiediamo al calcolatore umano di eseguire, e a ogni risorsa che gli mettiamo a disposizione.

Naturalmente, sono stato molto attento nello strutturare questo discorso su lavoranti, supervisori, caselle A, caselle P, e così via, in modo tale che queste risorse e funzioni fossero analoghe ai componenti e alle funzioni dei veri computer. In realtà, i veri computer hanno davvero caselle A, cioè registri chiamati accumulatori, e caselle P, cioè registri chiamati contatori di programma. I veri computer memorizzano davvero le loro istruzioni in registri numerati in modo sequenziale, e gestiscono il flusso di controllo più o meno nel modo che ho delineato. I computer moderni sono molto più complessi, sia nella loro struttura che nelle loro funzioni, di quanto appaia dal mio semplice schema di un tipo di computer, questo è certo. Ma i principi fondamentali di operazione che ho descritto forniscono un punto di partenza per capire anche i computer moderni.

Parlando di macchine di Turing, ho affermato che un linguaggio è un gioco giocato con un certo alfabeto e regolato da un insieme di regole di trasformazione, e che un computer è un'incarnazione di quelle regole. Questa affermazione si applica con ugual forza ai moderni computer digitali. Le procedure, come abbiamo visto, possono essere codificate come sequenze di bit il cui for-

mato è imposto dalle caratteristiche di progetto del computer cui sono destinate. Il vocabolario primitivo a disposizione di un programmatore è determinato dalle operazioni che esistono nel computer come operazioni interne. Per esempio, un particolare computer può avere la radice quadrata tra le sue operazioni primitive. Il linguaggio corrispondente a quel computer conterrà un codice d'operazione corrispondente a quella funzione, in pratica un verbo che significa «prendi la radice quadrata». Nel linguaggio determinato da un altro computer, che non abbia «prendi la radice quadrata» come operazione primitiva, la funzione radice quadrata può essere realizzata soltanto con una subroutine, in pratica un'espressione verbale composta da termini più elementari. Voglio di nuovo sottolineare che ogni computer determina un linguaggio, il suo linguaggio-macchina. L'alfabeto del linguaggio-macchina di tutti i computer moderni è l'insieme formato dai due simboli «0» e «1». Ma il loro vocabolario e le loro regole di trasformazione possono differire molto.

Ma allora, i programmatori di computer lavorano nel linguaggio della macchina che stanno istruendo? Vale a dire, quello che fanno è proprio comporre delle procedure complesse in forma di sequenze di bit? No. Abbiamo visto che è facile ed elementare tradurre l'espressione

(ADD 123)

nella riga di bit

00110101000000000000000001111011

Un computer è un superbo manipolatore di simboli. È relativamente facile progettare una procedura che converta un programma come quello della Tabella 3.2 nella riga di bit corrispondente. Anzi, procedure simili possono essere create anche per eseguire lavori di contabilità più prosaici, ma spesso tutt'altro che banali, come l'assegnazione dettagliata di posizioni di memoria. I programmi che fanno questo tipo di lavoro, che cioè convertono programmi scritti in una notazione come quella della Tabella 3.2 in linguaggio-macchina, sono chiamati *assemblatori*. Il linguaggio da cui convertono, cioè il linguaggio che traducono, è chiamato *linguaggio d'assemblaggio*.

I programmi scritti in assembly language sono, naturalmente, molto più facili da leggere, per noi, dei programmi scritti in linguaggio-macchina. Dal momento che gli assemblatori sono essi stessi dei programmi piuttosto complessi, sono scritti anch'essi in assembly language, cioè nel linguaggio che essi stessi traducono. In questo non c'è nulla di paradossale. Un assemblatore è un'incarnazione procedurale di regole ben determinate, che dicono come trasformare espressioni composte in un certo alfabeto in espressioni composte in un alfabeto molto più ristretto. Definisce perciò un linguaggio, il suo linguaggio d'assemblaggio. Poiché è progettato per tradurre qualunque testo lecito in linguaggio d'assemblaggio in un testo in linguaggio-macchina, non c'è nien-

te di misterioso nella sua capacità di tradurre anche il testo che costituisce la sua stessa definizione.

Ho detto in precedenza che un programma trasforma un computer in un altro computer. Un computer generale, non specializzato, in cui si carichi soltanto un programma per fare radici quadrate, diventa un computer specializzato e capace di calcolare soltanto radici quadrate. Un assemblatore, dunque, trasforma anch'esso il computer in cui è caricato. La trasformazione che induce ha importanti conseguenze: un programmatore che istruisce un computer usando solo un assemblatore non avrà mai bisogno di imparare il linguaggio determinato dal computer stesso, cioè il suo linguaggio-macchina. In un senso importante, egli non vede mai la macchina cui in realtà si sta rivolgendo; vede e lavora con un artefatto simbolico che, per lui, è la macchina.

Non c'è voluto molto tempo perché i programmatori si rendessero conto che la potenza di manipolazione dei simboli che il computer ha può essere usata per tradurre linguaggi di livello ancora «superiore». Consideriamo, per esempio, il seguente frammento di testo in linguaggio d'assemblaggio:

GET *a* ,

ADD *b* ,

STO *c* .

Anche soltanto in base alle poche nozioni presentate nelle pagine precedenti, sappiamo riconoscere che tale frammento deve essere una codificazione dell'espressione algebrica

$$c = a+b \ .$$

Chiaramente, basterà una procedura molto semplice per tradurre quest'ultima espressione nel frammento di linguaggio di assemblaggio mostrato sopra. Ora, senza soffermarsi troppo sui dettagli, si può correttamente dire che quanto segue è una codificazione della procedura per ottenere radici quadrate con cui ormai abbiamo tanta familiarità:

```
Procedura SQRT(n):
    Poni tolleranza = .001;
    Poni vecchia approssimazione = 0;
    Poni nuova approssimazione = 1;
    Finché | vecchia - nuova approssimazione | >, itera:
        vecchia approssimazione = nuova approssimazione;
        nuova approssimazione = (vecchia approssimazione **2 + n)/2* vecchia
        approssimazione;
    Fine;
    Risultato = nuova approssimazione;
Fine della procedura SQRT.
```

Tradurre questa procedura in linguaggio d'assemblaggio è relativamente semplice. Essenzialmente, è quello che abbiamo fatto in una delle pagine precedenti. È ben più difficile, tuttavia, scrivere una procedura che traduca in assembly language un qualunque programma scritto nel linguaggio di cui la procedura appena citata è un esempio. A prima vista, può sembrare che questa difficoltà risulti dalla ovvia complessità di quel linguaggio, rispetto alla evidente semplicità dei linguaggi d'assemblaggio. Ed è vero che il traduttore deve essere in grado di maneggiare una grande varietà di forme sintattiche, molte delle quali — specialmente poiché si presentano in combinazioni diverse — sono causa di problemi tecnici veramente difficili. Tuttavia, non sono questi i problemi più ardui.

Dobbiamo chiederci cosa significhi dire che la procedura che abbiamo mostrato è un «esempio» di un certo linguaggio. Di quale linguaggio? Il problema prioritario, e di gran lunga il più difficile, è proprio progettare il linguaggio «di livello superiore». Le regole di formazione e di trasformazione di un linguaggio (come di un gioco) sono, dopo tutto, un sistema tanto di vincoli, di proibizioni, quanto di permessi. Il fatto stesso che i «mattoni» fondamentali usati per comporre programmi in assembly language siano così profondamente elementari concede al programmatore un enorme numero di gradi di libertà. Un linguaggio d'assemblaggio per una particolare macchina è virtualmente una trasposizione mnemonica del suo linguaggio-macchina. Il progettista di un assembly language, perciò, non ha praticamente di fronte alcun problema di significato. Ma il progettista di un linguaggio di livello superiore deve decidere quali gradi di libertà permettere agli autori di programmi scritti in quel linguaggio. Ogni libertà che un autore esercita nel fare una scelta è usurpata al programmatore che dovrà usare il suo prodotto. Consideriamo un esempio molto semplice, e cioè l'espressione

$$d = a + b*c$$

(dove «$*$» è l'operatore di moltiplicazione). Un programmatore è libero di scrivere o il codice d'assemblaggio

GET *a,*

ADD *b,*

MPY *c,*

STO *d,*

che significa che *c* va moltiplicato per la somma di *a* e *b*, e il prodotto memorizzato in *d,* oppure

GET *b*,

MPY *c*,

ADD *a*,

STO *d*,

che significa che *a* va sommato al prodotto di *b* e *c*, e la somma va memorizzata in *d*. Ma l'espressione

$$d = a + b*c$$

nel linguaggio di livello superiore può avere soltanto una interpretazione. Quale questa debba essere è determinato dai dettagli del progetto del traduttore, e quindi dal progettista del linguaggio. Naturalmente, ciò non significa che l'espressività del linguaggio di livello superiore sia necessariamente limitata. Il linguaggio può, per esempio, permettere che le due diverse interpretazioni siano espresse, rispettivamente, da

$$d = (a+b) * c$$

e da

$$d = a + (b*c).$$

Benché questo esempio estremamente semplice non ne sia che un accenno, il problema non è stabilire quali procedure siano esprimibili in un linguaggio di livello superiore — per la maggior parte, questi linguaggi sono in realtà universali nel senso di Turing —, bensì quale sia lo stile di programmazione imposto dal linguaggio. Abraham Maslow, lo psicologo, disse un giorno che a una persona che ha solo un martello, tutto il mondo sembra un chiodo. Anche un linguaggio è un utensile, e anch'esso determina, in modo molto forte, il modo in cui chi lo usa vede il mondo.

Un programmatore solitario, che scrive un programma destinato a risolvere qualche problema peculiare della sua ricerca privata, non ha bisogno di preoccuparsi di come il suo programma potrebbe condizionare la visione del mondo di altre persone. I linguaggi di livello superiore, invece, sono destinati a essere proprio dei linguaggi pubblici, in due sensi importanti: sono destinati a essere usati da un gran numero di programmatori, per istruire le loro macchine; e sono anche destinati a fungere da linguaggi in cui persone diverse si comunicheranno le procedure che hanno composto. Infatti, il successo di un particolare linguaggio di programmazione di livello superiore si misura per lo più da quanto esso è diventato dominante sul mercato, per così dire, cioè dalla sua capacità di rendere obsoleti gli altri modi di esprimersi.

La storia recente delle scienze comportamentali ci ha mostrato quanto profondamente il successo della matematica applicata alla fisica abbia condizionato discipline ben lontane dalla fisica. Molti psicologi e sociologi, per esempio, hanno discusso per generazioni la loro materia in termini di equazioni differenziali e di statistica. Magari, hanno iniziato credendo che i metodi matematici adottati fossero semplicemente una notazione conveniente per descrivere i fenomeni con cui hanno a che fare. Ma, via via che hanno costruito sistemi concettuali sempre più vasti, a partire da componenti elementari originariamente presi a prestito da contesti estranei, e hanno dato dei nomi a questi sistemi e li hanno manipolati come elementi di sistemi di pensiero ancor più elaborati, quei sistemi stessi hanno smesso di essere soltanto dei modi di descrivere e sono diventati — come il martello di Maslow — determinanti per la loro visione del mondo. Il progetto di un linguaggio pubblico, dunque, è una impresa seria, densa di conseguenze e perciò carica di responsabilità straordinariamente pesanti. Ne riparlerò più avanti.

Per il momento, desidero sottolineare che un linguaggio per programmazione di livello superiore, come quello rappresentato dal frammento precedente, è in realtà un linguaggio formale. Il significato di espressioni scritte in quel linguaggio è determinato, naturalmente, dalle sue regole di trasformazione e queste, a loro volta, sono incarnate dalle procedure che lo traducono in linguaggio d'assemblaggio e, infine, in linguaggio-macchina. Se, perciò, si dovesse dire che cosa significa un programma scritto in un particolare linguaggio di livello superiore, bisognerebbe in ultima analisi indirizzarsi alla sua riduzione in linguaggio-macchina, e addirittura alla macchina che gli corrisponde. Si dovrebbe dire che «significa ciò che quella macchina fa con quel codice». Tuttavia, nella realtà non è così che si risponde a simili domande. Infatti, il traduttore è esso stesso un programma, e trasforma il computer in cui è caricato in un computer diverso, e cioè un computer per cui quel linguaggio è il linguaggio-macchina. C'è dunque un'equivalenza non solo tra linguaggi formali e giochi astratti, ma anche tra questi e le macchine calcolatrici. Per dirla in altre parole, esistono degli universi di discorso importanti in cui la distinzione tra linguaggio e sua incarnazione nella macchina scompare.

Temendo che questo possa sembrare un punto «solamente» filosofico e privo di conseguenze pratiche, devo immediatamente dire che non soltanto la maggior parte dei programmatori odierni pensano che il linguaggio che usano sia la loro macchina, proprio quasi in senso letterale, ma molti, forse i più, non sanno nulla del linguaggio-macchina del loro computer o del contenuto e della struttura dei traduttori che mediano tra loro e i loro computer. Questa osservazione non va intesa come una critica. Dopo tutto, un linguaggio formale di livello superiore è una macchina astratta. Non è una colpa il fatto di usare un linguaggio o una macchina più congeniale ai nostri scopi di qualche altra, semplicemente perché quest'ultima è in qualche modo più primitiva. Ma l'osservazione solleva una questione importante: se i programmatori odierni sono in larga misura ignari della struttura dettagliata delle macchine che usano, dei loro linguaggi e dei traduttori che manipolano i loro programmi, allora devono anche in larga misura ignorare molte delle tesi che ho sostenuto qui,

particolarmente quelle che riguardano l'universalità dei computer e la natura delle procedure efficaci. In che modo dunque avvertono, questi programmatori, la potenza del computer?

La loro convinzione che, per dir così, il computer possa fare qualunque cosa — vale a dire, l'intuizione corretta del fatto che i linguaggi a loro disposizione sono, in un certo senso non banale, universali — deriva in larga misura dall'impressione di poter programmare qualsiasi procedura che abbiano completamente capito. Questa impressione, a sua volta, è basata sulla loro esperienza della potenza delle subroutine e della riducibilità di processi decisionali complicati a gerarchie di scelte binarie (cioè, che si diramano lungo due possibili strade).

Una subroutine è, come ho detto, un programma di cui possiamo capire il comportamento di input/output senza dover necessariamente capire come faccia a trasformare i dati di input che gli diamo in quelli di output che ci restituisce. Un programmatore che abbia di fronte un sotto-problema difficile può sempre fare finta che esista una subroutine che risolve quel sotto-problema. Magari, la troverà da qualche parte in una biblioteca di programmi. Ma anche se non la trova, potrà dedicarcisi dopo aver elaborato la strategia per risolvere il suo problema principale. Se, per esempio, il problema principale è quello di scrivere un programma che permetta a un computer di giocare a scacchi, egli sicuramente avrà bisogno di una subroutine che produca una mossa, data una configurazione dei pezzi sulla scacchiera, e un'indicazione di chi è a muovere, se il bianco o il nero. L'input da dare a una simile subroutine deve essere una qualche rappresentazione dei pezzi sulla scacchiera e, mettiamo, uno «0» se la mossa è del bianco e un «1» se è del nero. L'output sarà una configurazione dei pezzi, data nella stessa rappresentazione usata per l'input. Quando ha deciso quali sono le rappresentazioni adatte — e già questo può essere un compito non banale — , il programmatore è libero di «usare» una subroutine che genera mosse per sviluppare ulteriormente il suo programma principale, come se tale subroutine esistesse. Se egli fa parte di un gruppo di programmatori, potrà magari descrivere il comportamento di input/output desiderato a un collega, il quale poi scriverà quella subroutine per lui. In ogni caso, egli può proseguire nella sua analisi strategica, e addirittura nella programmazione del problema principale, senza dover tenere sempre presenti dettagli di livello inferiore.

Ho già detto che, nei computer e nel calcolo, entità più grandi vengono costruite a partire da entità più piccole. In un certo senso, le subroutine ci permettono di dire anche l'opposto. Infatti, con il loro aiuto possiamo spezzare un problema molto vasto e complicato in un insieme di problemi più ristretti, ciascuno di questi in problemi ancora più ristretti, e così via. Se non fosse così, nessuno oserebbe iniziare imprese monumentali come la gestione computerizzata del traffico aereo, o la simulazione al computer del funzionamento di una grande azienda. Naturalmente, l'idea di risolvere un problema spezzandolo in problemi più piccoli, e poi risolvendo questi, è tutt'altro che nuova. Ma fino all'avvento del computer, gli enormi sistemi gerarchici per la soluzione dei problemi dovevano dipendere da esseri umani, come esecutori dei progetti elabo-

rati dagli strateghi dei sistemi. Gli esseri umani, a differenza dei computer, non sono totalmente affidabili come esecutori, neppure per le istruzioni più esplicite. Il senso di potenza del programmatore di computer deriva in larga misura dalla sua convinzione che le sue istruzioni verranno eseguite in modo incondizionato e che, data la sua capacità di costruire programmi arbitrariamente grandi, non ci sono limiti (per lo meno dal punto di vista delle dimensioni) ai problemi che egli sarà in grado di risolvere.

Sono finalmente in grado di fare qualche commento su un'impressione erronea che certamente è stata lasciata da qualcosa che ho scritto all'inizio del capitolo 2. Scrivevo che «le macchine fanno solo quello per cui sono state costruite; e lo fanno in modo esatto». Abbiamo visto ora come possiamo insegnare ai computer a fare qualcosa. Se prendiamo sul serio, come dovremmo fare, l'affermazione secondo cui un programma trasforma un computer in un altro, allora abbiamo anche visto come possiamo «fare» fare qualcosa a una macchina. Ma l'impressione che probabilmente ho generato è che i programmi debbano necessariamente avere la forma — magari sottilmente mascherata — di una sequenza di istruzioni da eseguire in modo più o meno sequenziale.

Ho già presentato prove a sostegno del contrario. Le quintuple che governano il comportamento di una macchina di Turing classica non devono essere in alcun ordine particolare, né vengono eseguite in modo sequenziale. Una tipica quintupla di quel tipo, messa in forma di regola di un gioco, dice:

> Se sei nello stato *u* e se il simbolo che stai scrutinando in questo momento è *x*, allora fai....

La macchina di Turing va letteralmente alla ricerca, sul suo nastro, della regola che si applica alla sua situazione. L'ordine in cui le regole sono scritte è del tutto irrilevante per il comportamento finale di input/output della macchina.

A questo proposito, ci sono due importanti osservazioni da fare. La prima riguarda la ricerca di una regola. La macchina di Turing va alla ricerca, in modo lineare, della regola da applicare. Dopo tutto, l'informazione sul suo nastro è memorizzata in ordine lineare. Ma nei computer moderni sono possibili altri regimi di ricerca. In particolare, si possono scrivere programmi la cui funzione è calcolare regole per la ricerca. Così, proprio come un programmatore può conoscere una procedura per calcolare le radici quadrate senza sapere la radice quadrata di alcun numero, e certamente non di tutti i numeri cui la sua procedura può venir applicata, egli potrà sapere come specificare la costruzione di regole di ricerca, senza conoscere alcuna delle regole che la sua specificazione potrà produrre. Un computer che obbedisca a regole che trova funzionando in questo regime di «ricerca calcolata» fa davvero soltanto ciò per cui è stato istruito. Ma il «fare» pratico è un passo (o in certi casi molti passi) più avanti rispetto al «fattore» originale, cioè al programmatore. Ritornerò più avanti sulla discussione di alcune implicazioni di questo problema.

La seconda osservazione riguarda il modo in cui regole vagamente simili alle quintuple delle macchine di Turing possono venir assegnate ai computer.

Certamente, delle regole possono essere introdotte sotto forma di programmi del tipo che ho mostrato. Ma dimentichiamoci per un attimo dei computer, e vediamo come una persona reagirebbe a una domanda del tipo «Quant'è l'area del piatto?». Per prima cosa, il nostro interlocutore supporrebbe che il piatto in questione è l'ultimo piatto di cui si è parlato. Cercherebbe perciò nella sua memoria la più recente menzione di un piatto. Se tale menzione fosse «L'area del piatto è tot», egli saprebbe subito rispondere. Altrimenti, ricercherebbe nella sua memoria affermazioni del tipo «L'area della tal cosa è...», cioè qualche conoscenza memorizzata su come si risponde a domande che riguardano l'area di qualcosa. Supponiamo che trovi la nozione «L'area di un cerchio è *pi greco* volte il quadrato del suo raggio». Potrebbe pensare: «Se il piatto ha un raggio, può darsi che sia un cerchio e in tal caso potrei applicare questa regola». Se venisse fuori che l'ultima citazione di un piatto era «Il raggio di un piatto è 10», il resto sarebbe ovvio. Quella che segue è la trascrizione di una «conversazione» con un computer condotta usando un terminale costituito da una macchina per scrivere elettrica (le righe tutte maiuscole sono le risposte del computer)[3].

L'area di un cerchio è pi greco *volte il raggio del cerchio al quadrato.*
L'area di una sfera è pi greco *volte il diametro della sfera al quadrato.*
L'area di un quadrato è il lato del quadrato al quadrato.
Il diametro di un oggetto è due volte il suo raggio.
Il raggio di una palla è 10.
Quant'è l'area della palla?
SUPPONGO CHE LA PALLA SIA UN CERCHIO; POSSO ANDARE AVANTI?
no.
SUPPONGO CHE LA PALLA SIA UNA SFERA; POSSO ANDARE AVANTI?
sì.
L'AREA È 1256. 6359.
Quant'è il diametro della palla?
È 20.
Il raggio di un piatto è 10.
Quant'è l'area del piatto?
SUPPONGO CHE IL PIATTO SIA UN CERCHIO; POSSO ANDARE AVANTI?
sì.
L'AREA È 324.15899.
Il raggio di un globo è 10.
Un globo è una sfera. Una sfera è un oggetto.
Quant'è l'area del globo?
È 1256.6359.

[3] Da J. Weizenbaum, *Contextual Understanding by Computers* («Comprensione contestuale da parte dei computer»), *Communications of the ACM*, vol. 10, 1967, nr. 8, pp. 474-480.

Il lato di un vassoio è 10.
Quant'è l'area del vassoio?
SUPPONGO CHE IL VASSOIO SIA UN QUADRATO; POSSO ANDARE AVANTI?
sì.
L'AREA È 100.
Quando dico piano intendo vassoio.
Quant'è 3 volte il lato del piano?
È 30.
Per favore, chiamalo *x*.
Supponiamo che f(z) *sia* z *al quadrato meno* z.
Calcola *f(x)*.
È 870.

Il programma che ha istruito il computer a fare questo è descrivibile, più che come un insieme di istruzioni del tipo «se questo è vero, fai quello», come una collezione di regole di evidenza. Usa queste regole per decidere in che modo le informazioni che gli vengono date si collegano alle domande che riceve. Il comportamento del computer perciò è governato tanto dalle informazioni che gli si danno, per esempio «Un globo è una sfera», quanto dalle sue istruzioni. La facile affermazione, spesso ripetuta, che un computer può fare soltanto ciò che gli si dice di fare, risulta — come la maggior parte di queste «verità» riguardanti argomenti complessi — essere a dir poco problematica. Ci sono molti modi per «dire» qualcosa a un computer.

L'idea che una persona possa scrivere un programma che incarni qualcosa che essa ha «capito completamente» è almeno altrettanto problematica. Capire qualcosa significa sempre capirlo a un certo livello. Un banditore d'asta usa alcuni strumenti matematici piuttosto sofisticati, di cui quasi certamente non capisce i fondamenti (né gliene importa granché). Chiunque lavori al cambio di valuta usa l'aritmetica, ma ben poche persone al mondo conoscono o si interessano dell'affascinante sistema di assiomi su cui l'aritmetica è basata. In realtà, tutti noi usiamo in continuazione delle subroutine di cui crediamo di conoscere il comportamento di input/output, ma dei cui dettagli non abbiamo bisogno di sapere molto, e ci pensiamo solo di rado. Capire qualcosa sufficientemente bene da poterlo programmare per un computer non significa capirlo fino nei suoi recessi più profondi. Nelle situazioni pratiche, questa comprensione ultima non ci può essere. La programmazione è piuttosto un test di comprensione. Da questo punto di vista, è come scrivere; spesso, quando pensiamo di aver capito qualcosa e proviamo a scriverne, proprio la nostra composizione ci rivela la nostra mancanza di comprensione. La nostra penna scrive la parola «poiché», poi si ferma. Pensavamo di aver capito il «perché» di qualcosa, ma scopriamo che non era vero. Iniziamo una frase con «ovviamente», e poi ci accorgiamo che quanto intendevamo scrivere non è affatto ovvio. Talvolta colleghiamo due frasi con la parola «perciò», soltanto per accorgerci che la nostra catena di ragionamenti è carente. La programmazione è così. Dopo tutto, è anch'essa un modo di scrivere. Ma nello scrivere ordinario talvolta na-

scondiamo la nostra mancanza di comprensione, le nostre carenze logiche, facendo ineffabilmente appello all'immensa flessibilità del linguaggio naturale, e alla sua intrinseca ambiguità. Proprio l'eloquenza permessa dal linguaggio naturale illumina talvolta le nostre parole e sembra (di sicuro falsamente) illuminare altrettanto brillantemente la nostra logica immeritevole. Un interprete di testi in linguaggio di programmazione, un computer, è immune dall'influenza seducente della pura eloquenza. E parole come «ovviamente» non trovano posto nel vocabolario primitivo di alcun computer. Un computer è un critico spietato. Perciò, l'affermazione secondo cui uno capisce una cosa sufficientemente bene da essere capace di programmarla è, prima di tutto, un'affermazione del fatto che la capisce in termini molto particolari. In ogni caso, può non essere altro che una presunzione, che magari sarà dimostrata falsa dall'esperienza.

L'altra faccia della medaglia è la credenza che non si possa programmare qualcosa a meno di capirla completamente. Qui si dimentica una verità, e cioè il fatto che il programmare (come, di nuovo, ogni forma di scrittura) è spesso una tecnica sperimentale. Si programma, proprio come si scrive, non perché si capisce, bensì per arrivare a capire. La programmazione è un'attività di progetto. Scrivere un programma significa emanare le leggi per un mondo che va prima creato nella fantasia. Soltanto molto raramente un progettista, che sia un architetto, un romanziere o chissà cos'altro, ha nella sua immaginazione un quadro così chiaro del mondo che creerà, da poterne scrivere le leggi senza tener conto delle critiche di quello stesso mondo. Ed è proprio questo che il computer fornisce.

Bisogna, naturalmente, imparare a usare le critiche. In molti casi, la critica del computer è radicale e non può essere ignorata; un programma non funziona affatto, oppure produce risultati chiaramente sbagliati. Allora, è fuori discussione che qualcosa va rifatto. I computer sono maledettamente efficienti nell'inciampare su errori puramente tecnici (cioè linguistici, di programmazione), e nel farlo in un modo che occulta la vera sede del problema, per esempio qual è la parentesi messa al posto sbagliato. Infatti, molti programmatori di professione credono che la loro arte sia difficile perché devono usare un linguaggio che ha regole sintattiche molto rigide. C'è perciò una continua richiesta di sistemi di programmazione in linguaggio naturale, per esempio in inglese. I programmatori che partecipano a questo coro probabilmente non hanno mai affrontato un problema veramente difficile, e perciò non hanno mai sentito il bisogno di una critica veramente profonda da parte del computer. È vero che per poter scrivere bisogna prima padroneggiare le regole sintattiche del linguaggio che si è scelto. Ma, oltre a questo, bisogna anche aver pensato bene a ciò che si vuole scrivere. La critica letteraria non consiste nel mettere in evidenza gli errori di ortografia e le violazioni tecniche di regole grammaticali. Riguarda invece la sostanza delle idee. Il critico letterario deve sapere molto. Una vera ragione per cui programmare è difficile è che, nella maggior parte dei casi, il computer non sa nulla di quegli aspetti del mondo reale che il programma intende trattare. È in una posizione analoga a quella, mettiamo, di un insegnante di grammatica che si ritrova, con un manuale di pilotaggio, a dover guidare un aeroplano in volo. Può darsi che sia molto bravo a trovare

errori grammaticali, ma quanto a pilotare, non ne sa nulla. C'è quindi da sperare che il manuale spieghi tutto in termini abbastanza primitivi da permettere anche a un maestro di scuola, abituato a stare a terra, di capirli. Quanto più facile sarebbe scrivere un manuale di pilotaggio per uno studioso di aerodinamica che, benché ancora incapace di pilotare un aereo, conosce la teoria del volo! È in realtà molto difficile spiegare qualsiasi cosa usando un vocabolario così primitivo da non avere niente a che fare con ciò che si vuole spiegare. Eppure, questo è proprio ciò che la maggior parte dei programmi tenta di fare. Le difficoltà che sorgono non hanno le loro radici nelle rigidità sintattiche più di quanto, mettiamo, la difficoltà di scrivere un buon sonetto dipenda dalla rigida forma richiesta da quel tipo di poesia. Per scrivere un buon sonetto, o un buon programma, uno deve sapere che cosa intende dire. E se il suo critico condivide la sua base di conoscenze, questo aiuta moltissimo.

Abbiamo visto un programma che consiste, in parte, in affermazioni date in linguaggio naturale. Esse possono venir fornite al programma a tempi diversi. L'insieme delle affermazioni può perciò crescere fino a diventare molto grande. In un certo senso, è più semplice programmare in questo stile che in quelli, più ortodossi, di uso comune. Ed è possibile, usando simili tecniche di programmazione, dotare un computer di una base di conoscenze. Un simile sistema di programmazione dimostra così che è possibile sfuggire al rigore formale normalmente associato al lavoro di programmazione.

Ma non dobbiamo dare a questo troppa importanza. Un elenco di formule di fisica utili può permettere a uno studente pigro di risolvere molti problemi da libro di testo, e tuttavia non gli darà una comprensione della fisica, una teoria su cui pensare. Allo stesso modo, un insieme di informazioni del tipo che abbiamo mostrato può permettere a un computer di risolvere certi problemi. Ma questo non vuol dire che il programmatore o il computer padroneggino una teoria, o che alcuno dei due abbia capito altro se non il modo di usare un gruppo di «fatti» per arrivare a certe conclusioni. Una prestazione del computer che ha successo viene spesso presa come dimostrazione che esso, o il suo programmatore, capisce la teoria in questione. Tale deduzione non è necessariamente vera e, il più delle volte, è del tutto fallace. La relazione tra comprensione e scrittura rimane così altrettanto problematica per la programmazione al computer quanto lo è sempre stata per la scrittura in qualsiasi altra forma.

4
LA SCIENZA E IL PROGRAMMATORE COATTO

C'è una differenza sostanziale tra le macchine dotate di un'incarnazione fisica, il cui fine ultimo è trasmettere energia o distribuire potenza, e le macchine astratte, cioè quelle che esistono soltanto come idee. Le leggi incarnate dalle prime devono essere un sottoinsieme delle leggi che governano il mondo reale. Le leggi che governano il comportamento delle macchine astratte non sono necessariamente vincolate da questa condizione. È possibile, per esempio, progettare una macchina astratta al cui interno i segnali si propaghino tra i diversi componenti a velocità superiore a quella della luce, violando con ciò chiaramente una legge fisica. Il fatto che una simile macchina non possa essere costruita non proibisce l'esplorazione del suo comportamento. Può essere oggetto di riflessione, e può perfino venire simulata al computer (infatti, il Centro Ricerche Educative del MIT di Cambridge, USA, ha prodotto dei filmati generati dal computer che permettono agli spettatori di osservare un mondo in cui i veicoli possono viaggiare a velocità fisicamente impossibili). La fantasia umana deve essere capace di trascendere le limitazioni della legge fisica, se vuole poter concepire quella stessa legge.

Il computer, naturalmente, è una macchina dotata di un'incarnazione fisica e, come tale, non può violare la legge naturale. Ma per caratterizzarlo completamente non bastano solo le sue interazioni manifeste col mondo reale. Elettroni scorrono al suo interno, nastri magnetici si muovono e spie si accendono in stretta osservanza delle leggi naturali, questo è certo. E il percorso dei fiumi di elettroni al suo interno è determinato dall'apertura e chiusura dei «cancelli», vale a dire da eventi fisici. Ma il gioco cui il computer partecipa è regolato da sistemi di idee il cui dominio è limitato soltanto dalle frontiere della fantasia umana. I vincoli di natura fisica sugli eventi elettronici e meccanici interni al computer non hanno importanza per quel gioco, non più di quanto importi il modo in cui il giocatore di scacchi solleva il suo alfiere, o quanto rapidamente lo sposta da una posizione all'altra sulla scacchiera.

Un computer che lavori sotto il controllo di un programma memorizzato è quindi altrettanto avulso dal mondo reale di qualsiasi gioco astratto. La scacchiera, i 32 pezzi e le regole degli scacchi costituiscono un mondo del tutto separato da qualsiasi altro mondo. Lo stesso vale per un sistema di computer

con il suo manuale operativo. Il giocatore di scacchi che ha fatto una mossa sbagliata non può giustificare il suo errore invocando qualche fatto empirico esterno che non poteva sapere ma che, se l'avesse saputo, lo avrebbe portato a fare una mossa migliore. Allo stesso modo, un programmatore il cui programma si comporti diversamente dalle intenzioni non può cercare lo sbaglio se non nel gioco che egli stesso ha creato. Può darsi che abbia sbagliato a leggere il manuale del suo sistema, o che abbia in qualche altro modo frainteso l'ambiente-computer in cui si sta muovendo, proprio come il giocatore di scacchi principiante può fraintendere, mettiamo, le regole sull'arrocco. Ma nessun dato esistente nel mondo esterno al sistema di computer che sta usando può essere rilevante per il comportamento del mondo che egli ha creato. Se un computer non si comporta esattamente come inteso dal suo programmatore, questo non può neanche essere attribuito unicamente a qualche limitazione peculiare di quel computer. In realtà, ogni computer per uso generale, non specializzato, è una sorta di macchina universale che può fare, in linea di principio, ciò che qualunque altro computer per uso generale può fare. Da questo importante punto di vista, un qualsiasi computer per uso generale non ha alcuna limitazione peculiare. Il computer, dunque, è un campo di gioco su cui si può giocare qualunque gioco immaginabile. Si possono creare mondi in cui non c'è gravità, o in cui due corpi si attraggono non secondo la legge newtoniana dell'inverso del quadrato della distanza, bensì secondo una formula che va come l'inverso del cubo (o dell'ennesima potenza) della distanza fra loro, o ancora mondi in cui il tempo danza avanti e indietro obbedendo a una coreografia complessa a piacere. Si possono creare società nella cui economia i prezzi crescono quando i beni di consumo sono abbondanti, e diminuiscono quando diventano scarsi, e in cui soltanto le unioni omosessuali generano prole. In breve, si possono con leggerezza scrivere e rappresentare commedie in un teatro ove non ci sono limitazioni. E, cosa più importante, è necessario sapere soltanto ciò che può essere dedotto direttamente dal proprio manuale del sistema, o costruito dalla propria fantasia.

Per contrasto, consideriamo l'atto della progettazione di un circuito per un computer. In precedenza, ho mostrato il diagramma del circuito per un sommatore a un bit (Figura 3.5, p. 87). È possibile costruire il sommatore corrispondente per una macchina il cui ciclo temporale fondamentale sia, diciamo, un microsecondo, in altre parole una macchina che ha un milione di coppie «intervallo di riposo/intervallo attivo» ogni secondo. Supponiamo che un simile sommatore venga costruito, e risulti funzionare correttamente. Mettiamo che il suo progettista lo installi, in seguito, su una macchina che lavora dieci volte più velocemente della prima (che cioè, per così dire, si accende e spegne dieci milioni di volte al secondo), e supponiamo che in questo nuovo ambiente il sommatore non funzioni più. Quale può essere il problema? Cosa più importante, quali sono le fonti di informazione cui fare ricorso per diagnosticare correttamente la disfunzione? Chiaramente, le regole del gioco stabilite dalle equazioni astratte che governano il comportamento dei cancelli AND, OR e NOT e dal diagramma di circuito del sommatore non sono sufficienti. Niente è stato cambiato rispetto al progetto astratto del sommatore. Per ipotesi, l'u-

nica cosa cambiata è il ritmo a cui esso viene fatto funzionare. Dev'essere perciò il nuovo ritmo a causare la disfunzione. Per scoprire come si fa a riparare il sommatore, o anzi per sapere se è fisicamente possibile far funzionare un qualsiasi apparecchio a un ritmo così elevato, è necessario applicare qualche opportuna nozione del mondo fisico.

Questo esempio, benché semplicistico, non è campato in aria, e vale la pena di esaminarlo più a fondo. Quasi tutti sanno che le corde di un violino (così come altri oggetti) hanno frequenze di vibrazione naturali (o risonanti). La corda del *la*, per esempio, vibra naturalmente circa 435 volte al secondo. Se due corde del *la* sono vicine fra loro, e una di esse viene fatta vibrare alla sua frequenza naturale, ad esempio pizzicandola, l'altra corda entrerà anch'essa in vibrazione (naturalmente, con la stessa frequenza). Se, però, la prima corda viene costretta a vibrare a una frequenza diversa da quella risonante, allora, nonostante entrambe siano corde di *la*, la seconda non vibrerà. In pratica, le corde sono antenne trasmittenti e riceventi alla loro frequenza acustica naturale. Tenendo presente questo, supponiamo che due bambini che vivono ai lati opposti di una strada costruiscano un sistema di segnalazione consistente in due corde parallele che passano sopra la strada, collegando le loro case. Per segnalare l'arrivo di papà, mettiamo, una delle corde viene fatta vibrare, e lo stesso vale per l'altra all'arrivo di mamma. Se, però, una corda viene fatta vibrare alla sua frequenza naturale, l'altra corda sarà anch'essa indotta a vibrare. Il sistema, perciò, non funzionerà come voluto: sarà incapace di annunciare l'arrivo di un solo genitore.

Da questo punto di vista, i circuiti elettrici sono abbastanza simili ai sistemi acustici, anche se le frequenze loro associate sono molto più alte. In particolare, anche i circuiti hanno frequenze naturali (o risonanti). Anzi, molti di loro possono essere accordati in modo da risuonare a frequenze molto precise, per esempio se si vuole farli funzionare da antenne per la trasmissione di programmi radio. Gli apparecchi radiofonici domestici sono dotati di circuiti simili, che fungono da antenne riceventi. Il problema del nostro sommatore potrebbe essere che, anche se nessuno dei suoi circuiti risuona, trasmettendo e ricevendo così segnali dagli altri circuiti, quando lo si fa lavorare a una frequenza di un milione di impulsi al secondo, i suoi circuiti diventano antenne trasmettenti e riceventi alla frequenza superiore. In conseguenza, possono confondere l'informazione che devono manipolare.

Gli ingegneri specializzati, naturalmente, conoscono bene la teoria da cui questi fenomeni possono venir dedotti. Altri possono identificare il problema sulla base della loro esperienza. Ma un ingegnere che non abbia né la teoria né l'esperienza non sarà mai in grado di scoprire la causa del problema usando soltanto la logica deduttiva. Potrà riparare il circuito soltanto per combinazione, a forza di tentativi, oppure ricevendo aiuto esterno, oppure ripetendo in pratica tutto il lavoro scientifico creativo di un uomo come Heinrich Hertz, lo scopritore del fenomeno della radiazione elettromagnetica cui stiamo alludendo. Ma il lavoro di Hertz lo chiamiamo «creativo» proprio perché implica l'osservazione di fenomeni non citati in alcun manuale esistente, e perché rica-

va leggi generali dai particolari osservati. Il che è proprio l'opposto della deduzione.

Un ingegnere è inestricabilmente invischiato nel mondo materiale. La sua creatività è limitata dalle leggi di quel mondo; egli può fare, in ultima analisi, solo ciò che è lecito fare. Ma è costretto a esercitare la sua attività in un castello kafkiano da cui non c'è uscita, nemmeno in linea di principio. Perché egli non può conoscere la mappa che determina quali stanze ci sono nel mondo, quali porte le collegano e come si fa ad aprirle. Quando un apparecchio che ha progettato non funziona, l'ingegnere perciò non può sempre sapere (o dire, soltanto in base al ragionamento) se si trovi nell'anticamera del successo, dove solo la sua miopia gli impedisce di aprire le porte, o piuttosto in uno sgabuzzino privo d'uscite. Allora deve rivolgersi ad altri — i suoi insegnanti, i colleghi, i libri — perché gli dicano, o almeno gli suggeriscano, una formula che costringa il guardiano indifferente (la natura) a guidarlo fuori dal labirinto.

Il programmatore di computer, però, è creatore di universi in cui egli solo è legislatore. Lo stesso, naturalmente, vale per il progettista di qualunque gioco. Ma universi di complessità virtualmente illimitata possono venir creati sotto forma di programmi per il computer. Inoltre, e questo è un punto cruciale, sistemi formulati ed elaborati in quel modo *recitano apertamente* i loro copioni programmati. Obbediscono diligentemente alle loro leggi, e mostrano con vivacità il loro comportamento obbediente. Nessuno sceneggiatore, nessun regista, nessun imperatore per quanto potente ha mai esercitato un'autorità così assoluta da organizzare un palcoscenico o un campo di battaglia e da comandare attori o truppe tanto fedelmente dediti al loro dovere.

Ci sarebbe di che stupirsi se l'osservazione di Lord Acton secondo cui il potere corrompe non si applicasse in un ambiente in cui è così facile raggiungere l'onnipotenza. E in effetti, si applica. E la corruzione evocata dall'onnipotenza del programmatore di computer si manifesta in una forma che è istruttiva in un dominio ben più vasto del ristretto settore dei computer. Per capirla, dovremo interessarci per un momento di una disfunzione mentale che, benché sia in realtà molto antica, sembra essere stata trasformata dal computer in una sindrome di tipo nuovo: la coazione a programmare.

Dovunque sono stati creati centri di calcolo, vale a dire in innumerevoli località degli Stati Uniti, così come virtualmente in tutte le altre regioni industriali del mondo, è possibile vedere giovani brillanti dall'aspetto slavato, spesso con gli occhi infossati e luccicanti, seduti alle tastiere del computer, le braccia in tensione nell'attesa di scatenare le dita, già pronte a colpire, sui pulsanti e sui tasti su cui la loro attenzione sembra essere concentrata come quella del giocatore d'azzardo sul rotolare del dado.

Quando non sono in questa *trance*, seggono spesso a tavoli inondati di stampati del computer, su cui stanno chini come studiosi posseduti dal demonio su un testo cabalistico. Lavorano fin quasi a crollare, venti, trenta ore per volta. Il cibo, se se ne ricordano, se lo fanno portare: caffè, coca-cola, panini. Se possibile, dormono su brande vicine al computer. Ma soltanto poche ore, poi di nuovo al terminale o agli stampati. I vestiti rattoppati, le facce non lavate e non rasate, i capelli spettinati dimostrano quanto si disinteressino del

loro corpo e del mondo che li circonda. Esistono, almeno nei periodi in cui lavorano così, soltanto attraverso e per il computer. Questi sono i pazzi del computer, i programmatori coatti. Sono un fenomeno internazionale.

Come si fa a distinguere il programmatore coatto da un programmatore di professione semplicemente volenteroso, e che lavora sodo? In primo luogo, dal fatto che il normale programmatore di professione si dedica al problema da risolvere, mentre il forzato vede il problema principalmente come un'occasione per interagire col computer. Il programmatore normale di solito discuterà i suoi problemi, tecnici e di sostanza, con altri. In genere, farà un lungo lavoro preparatorio, scrivendo o compilando diagrammi di flusso, prima di cominciare a lavorare col calcolatore. Le sue sessioni di lavoro col calcolatore potranno essere relativamente brevi. Potrà addirittura darsi che egli lasci ad altri il lavoro al terminale. Egli sviluppa il suo programma in modo lento e sistematico. Quando qualcosa non funziona, può passare una notevole quantità di tempo lontano dal computer, sviluppando attente ipotesi per spiegare la disfunzione e progettando esperimenti cruciali per metterle alla prova. Di nuovo, può darsi che lasci ad altri l'esecuzione pratica sul computer. È capace, nell'attesa dei risultati del computer, di dedicarsi ad altri aspetti del suo lavoro, come la documentazione di ciò che ha già fatto. Quando finalmente ha composto il programma che si era prefisso di fare, sa compilarne una descrizione intelligente, per poi dedicarsi ad altre cose. Il professionista considera la programmazione come un mezzo per uno scopo, non come un fine in se stesso. La sua soddisfazione deriva dall'aver risolto un problema di sostanza, non dall'aver piegato un computer alla sua volontà.

Il programmatore coatto è di solito un tecnico superbo, e in più uno che conosce tutti i dettagli del computer su cui lavora: il suo equipaggiamento periferico, il sistema operativo ecc. Spesso egli è sopportato, nel centro di calcolo, per la sua conoscenza del sistema e perché sa scrivere velocemente piccoli programmi di sottosistema («velocemente» significa, diciamo, in un paio di sessioni di venti ore ciascuna). Dopo un po' di tempo, può darsi in realtà che il centro usi un buon numero di suoi programmi. Ma poiché è impossibile convincere il coatto a fare altro che non sia il programmare, egli non documenterà quasi mai i suoi programmi, una volta finito di lavorarci sopra. Un centro può perciò finire per dipendere da lui per insegnare a usare, e aggiornare, i programmi che ha scritto e di cui soltanto lui capisce la struttura, ammesso che la capisca. La sua posizione è piuttosto simile a quella di un impiegato di banca che non fa molto per la banca, ma che viene tenuto perché soltanto lui sa la combinazione della cassaforte. Il suo interesse principale non sono, comunque, i programmi piccoli, bensì dei sistemi di programmi molto grandi e molto ambiziosi. Di solito i sistemi che si mette a costruire, e su cui lavora febbrilmente magari per un mese o due o tre, hanno obiettivi grandiosi, ma delineati in modo molto impreciso. Qualche esempio di queste ambizioni: nuovi linguaggi per facilitare la comunicazione uomo-macchina; un sistema generale cui si possa insegnare a giocare qualsiasi gioco da tavolo; un sistema che semplifichi per gli esperti di computer la scrittura di super-sistemi (quest'ultimo è il preferito). È tipico di molti progetti simili il fatto che il programmatore coatto può rima-

nere a lungo convinto che essi richiedano esclusivamente conoscenze di computer, programmazione ecc. E di tali conoscenze lui, naturalmente, ne ha in abbondanza. Anzi, il punto in cui tali lavori vengono spesso abbandonati è precisamente quando cessano di essere puramente incestuosi, vale a dire quando il lavoro di programmazione dovrebbe essere interrotto per permettere l'acquisizione di conoscenze dal mondo esterno al computer.

A differenza del professionista, il coatto non sa dedicarsi ad altro, neppure a mansioni strettamente legate al suo programma, nei periodi in cui non sta proprio facendo funzionare il computer. A stento sopporta di stare lontano dalla macchina. Ma quando, nonostante tutto, le circostanze lo separano da essa, ha almeno con sé gli stampati del computer. Li studia, ne parla con chiunque lo stia a sentire (benché, naturalmente, nessun altro possa capirli). Anzi, quando è vittima di un attacco non sa parlare d'altro che del suo programma. Ma gli unici momenti in cui è, per così dire, felice sono quelli in cui è alla *console*. Allora non parla con altri che con il computer. Vedremo presto di che cosa parlano.

Il programmatore coatto passa tutto il tempo che può a lavorare su uno dei suoi grandi progetti. Lui non usa la parola «lavorare»; chiama quello che fa *hacking*. Secondo il vocabolario, «to hack» significa «tagliare irregolarmente, in modo maldestro o senza uno scopo preciso; con (o come se lo si facesse con) ripetuti colpi di uno strumento affilato». Ho già detto che il programmatore coatto, o hacker come gli piace chiamarsi, è di solito un ottimo tecnico. Non sembra perciò che sia «maldestro» come vorrebbe la definizione. Ma la definizione è profondamente azzeccata nel dire che lo hacker è «privo di uno scopo preciso»: non sa prefiggersi un obiettivo a lunga scadenza chiaramente definito, poiché ha soltanto tecnica, non conoscenze. Non c'è niente che sappia analizzare o sintetizzare; in breve, non ha niente su cui fare teorie. La sua bravura è perciò priva di scopo, e addirittura incorporea. È semplicemente scollegata da qualunque cosa che non sia lo strumento su cui si eserciterà. La sua bravura è come quella dell'amanuense che, benché analfabeta, sia un calligrafo di prima qualità. I suoi progetti grandiosi devono perciò necessariamente avere il carattere di illusioni, anzi di illusioni di *grandeur*. Lui costruirà l'unico grande sistema in cui tutti gli altri esperti presto scriveranno i loro.

(Va detto che non tutti gli hacker sono patologicamente dei programmatori coatti. Anzi, se non fosse per la fatica spesso altamente creativa di gente che è fiera di farsi chiamare hacker, oggi esisterebbero pochi degli odierni sofisticati sistemi di condivisione del tempo-macchina [*time-sharing*], traduttori di linguaggio del computer, sistemi grafici computerizzati).

I sistemi di programmazione possono, naturalmente, venir costruiti senza progetto e senza conoscenza (per non parlare di comprensione) dei profondi problemi strutturali coinvolti, proprio come è possibile «ritagliare» e mettere insieme case, città, sistemi di dighe o politiche economiche nazionali. Quando però un sistema costruito in questo modo inizia a diventare grande, diventa anche sempre più instabile. Quando una delle sue sottostrutture si mette a funzionare in modo imprevisto, lo si può rattoppare fino a far scomparire quel problema evidente. Ma dal momento che non esiste una teoria generale del-

l'intero sistema, il sistema stesso può soltanto essere un'aggregazione più o meno caotica di sottosistemi la cui reciproca influenza può essere scoperta soltanto volta per volta, e in modo empirico. Lo hacker passa parte del suo tempo alla console, accumulando nuovi sottosistemi sulla struttura che ha già costruito — li chiama «nuove opzioni» —, e il resto del suo tempo in tentativi di spiegare perché le sottostrutture già introdotte non funzionino. È di questo che parla con il computer.

La situazione psicologica in cui il programmatore coatto si trova mentre è così impegnato è fortemente determinata da due fatti apparentemente in opposizione: primo, egli sa che può far fare al computer tutto quello che vuole; secondo, il computer gli porta continuamente delle prove innegabili del suo fallimento. Lo respinge. A questo non c'è modo di sfuggire. L'ingegnere può rassegnarsi di fronte alla verità che ci sono alcune cose che lui non sa. Ma il programmatore si muove in un mondo interamente creato da lui. Il computer sfida la sua potenza, non le sue conoscenze.

Infatti, l'eccitazione del programmatore coatto sale al suo picco massimo e più febbrile quando egli è alla caccia di un errore dei più ostinati, quando tutto dovrebbe funzionare eppure il computer lo respinge, funzionando male in un gran numero di punti, apparentemente non correlati fra loro. È allora che il sistema che il programmatore stesso ha creato sembra dimostrare di aver preso vita per conto suo e, certamente, di essere sfuggito al controllo del creatore. Ed è qui che l'idea che al computer si possa «far fare qualsiasi cosa» diventa più pertinente e più seriamente basata sulla realtà. Perché, in tali circostanze, l'artefatto che non funziona è, in realtà, la creazione propria del programmatore. Il malfunzionamento, come abbiamo già detto, non può che essere la conseguenza di ciò che il programmatore stesso ha fatto. E questo egli può presumibilmente riuscire a capirlo, a disfarlo e a rifarlo meglio. Ed è così che il suo umore e la sua attività impazziscono quando crede di aver finalmente scoperto la ragione del problema. Se per caso il tempo a sua disposizione alla *console* è, a quel punto, quasi terminato, egli sarà capace di correre enormi rischi, facendo dei cambiamenti sostanziali al programma, uno dopo l'altro, in pochi minuti o anche in pochi secondi, senza scrivere un appunto di quello che fa, implorando «ancora un altro minuto!». In queste circostanze, egli può distruggere rapidamente e pressoché irreversibilmente settimane e settimane del suo stesso lavoro. Ma se, invece, gli capita di trovare un errore sottilmente nascosto, che davvero rende ragione di gran parte del malfunzionamento del programma, la sua gioia sarà illimitata. È davvero un'emozione forte vedere un programma, fino a un momento fa moribondo, tornare improvvisamente alla vita. Quando un errore profondo viene scoperto ed eliminato, tante diverse parti del programma, che fino ad allora non avevano prodotto altro che risultati incomprensibili, improvvisamente si comportano come si deve, producendo esattamente i risultati attesi. Chi l'ha trovato ha una buona ragione per essere compiaciuto e, se l'errore era davvero profondo, anche fiero della propria bravura.

Ma l'orgoglio e la soddisfazione del programmatore coatto durano molto poco. Il suo successo consiste nell'aver dimostrato al computer chi è il padro-

ne. E avendo dimostrato che può fargli fare tutto questo, immediatamente si dedica a fargli fare ancora di più. Così, il ciclo ricomincia. Egli inizia a «migliorare» il sistema, mettiamo, facendolo girare più velocemente, o aggiungendogli «nuove opzioni», o migliorando il sistema di input e output dei dati. L'attività di modifica del programma esistente fa invariabilmente sì che qualcuna delle sue sottostrutture crolli; esse costituiscono, dopo tutto, una collezione amorfa di processi le cui mutue interazioni sono in pratica casuali. Quello che apparentemente è uno sforzo del programmatore per migliorare e promuovere la sua creazione è in realtà un assalto contro di essa, assalto la cui unica conseguenza può essere rinnovare la sua lotta contro il computer. Se gli si dovesse impedire di sabotare in questo modo il suo stesso lavoro (mettiamo, per decisione amministrativa), egli diventerà chiaramente depresso, comincerà a mettere il broncio, non mostrerà alcun interesse per ciò che lo circonda ecc. Soltanto una nuova occasione per calcolare potrà rimetterlo in sesto.

Va sottolineato che il ritratto che ho disegnato può essere riconosciuto immediatamente nei centri di calcolo di tutto il mondo. Esso rappresenta una patologia psichica di gran lunga meno ambigua rispetto, mettiamo, alle forme più lievi di schizofrenia o paranoia. Allo stesso tempo, rappresenta una forma di disordine mentale estremamente sviluppata, che affligge gran parte della nostra società.

Come va interpretata questa coazione? In primo luogo, bisogna rendersi conto che è davvero una coazione. Normalmente, il desiderio di soddisfazione porta a comportamenti che hanno un sostrato di discernimento e spontaneità. La realizzazione di questo desiderio procura soddisfazione. Il programmatore coatto è inibito; c'è poca spontaneità nel suo modo di comportarsi; e non prova piacere nella realizzazione di quelli che sono, ufficialmente, i suoi desideri. Dal computer cerca rassicurazioni, non piacere. Il parallelo più vicino che si può trovare a questo tipo di psicopatologia è l'incessante spinta alla rassicurazione che caratterizza la vita del giocatore coatto (anche qui, il piacere è assente).

Il giocatore coatto va anch'esso chiaramente distinto dal giocatore di professione. Quest'ultimo, da un importante punto di vista, non è affatto un giocatore d'azzardo (possiamo lasciare da parte il baro, e il «finto pollo» che si fa spennare, poiché nemmeno loro certamente sono giocatori d'azzardo). Il cosiddetto giocatore di professione è in realtà un esperto di statistica applicata, e magari anche di psicologia. Le sue entrate non dipendono quasi per nulla dalla fortuna. Egli conosce la teoria della probabilità applicata, e la usa per calcolare le quote, per poi giocarle in combinazioni tali da permettergli di prevedere il suo guadagno su un periodo, mettiamo, di un anno, con certezza quasi matematica. Questo non è gioco d'azzardo. Ci sono poi persone che giocano senza essere né giocatori di professione, né coatti. Per il giocatore coatto, il gioco è tutto. Vincere è perfino meno importante che giocare. È felice, se così si può dire, soltanto al tavolo da gioco.

Chiunque abbia lavorato in un centro di calcolo, o in un casinò che chiuda la notte, riconoscerà la scena descritta da Dostoevskij, anch'egli giocatore appassionato, ne *Il giocatore d'azzardo*:

> Verso le undici, rimangono al tavolo della roulette soltanto quei giocatori disperati, i veri giocatori d'azzardo, per cui non esiste altro che il tavolo, [...] che non sanno nulla di ciò che accade intorno a loro e non hanno il minimo interesse per faccende esterne alla sala della roulette, ma non fanno che giocare dalla mattina alla sera, e sarebbero lieti di giocare tutta la notte, se gli fosse permesso. Queste persone sono sempre seccate quando arriva la mezzanotte e devono andare a casa, perché il banco della roulette ha chiuso. E quando il capo-croupier, a mezzanotte in punto, appena prima della chiusura, chiama «Gli ultimi tre giri, signori!» questi uomini sono pronti a puntare tutto quello che hanno in tasca in quegli ultimi tre giri, ed è certo che è allora che perderanno di più [1].

Questa di Dostoevskij potrebbe anche essere stata la descrizione di una sala di computer.

La letteratura medica sul gioco coatto riguarda per lo più la psicogenesi di quella coazione, e la affronta quasi interamente da un punto di vista psicoanalitico. Ma non c'è bisogno di ricapitolare la tesi psicoanalitica, per quello che è il mio obiettivo qui. Basti dire che gli psicoanalisti, a cominciare da Freud, hanno visto megalomania e deliri di onnipotenza come ingredienti principali della vita psichica del giocatore coatto. Non dobbiamo necessariamente accettare, o rifiutare, le prove psicoanalitiche dell'origine di questo delirio — per esempio, la tesi che sia radicato in conflitti edipici non risolti, che portano a un desiderio di sopraffare il padre; desiderio che, a sua volta, motiva inconsciamente a fallire — per poterci unire agli psicoanalisti e ai romanzieri come Dostoevskij nel riconoscere il ruolo centrale che le fantasie megalomaniache di onnipotenza hanno nel gioco coatto.

Il giocatore d'azzardo, secondo lo psicoanalista Edmund Bergler, ha tre convinzioni principali: primo, egli è soggettivamente sicuro di vincere; secondo, ha una fede illimitata nella propria bravura; terzo, sa che anche la vita non è che un gioco d'azzardo [2].

Che motivo ci può mai essere per ritenere che si vincerà a un gioco basato sul puro caso? Sapere che il lancio di un paio di dadi o l'estrazione di una carta è un evento puramente casuale equivale a sapere che è impossibile fare alcunché per condizionarne il risultato. Ma proprio qui sta il trucco! Il giocatore coatto crede di avere il controllo di un mondo magico cui solo pochi uomini hanno accesso. «Egli crede» scrive Bergler «che il destino l'abbia scelto ... e che comunichi con lui usando piccoli segni che indicano approvazione o diniego» [3]. Il giocatore è lo scienziato di questo mondo magico. È l'interprete dei segni che il Fato gli comunica, proprio come lo scienziato del mondo reale è l'interprete dei segnali che la natura comunica a chiunque si preoccupi di farvi attenzione. E come lo scienziato naturale, il giocatore coatto ha sempre pronta un'ipotesi che rende ragione di quasi tutti i segnali che ha osservato finora,

[1] F. Dostoevskij, *The Gambler* («Il giocatore»), cit. in E. Bergler, *The Psychology of Gambling* («La psicologia del gioco d'azzardo») Hill and Wang, New York 1957, p. 33.

[2] Bergler, *op. cit.*, p. 9.

[3] *Ivi*, p. 230.

un'ipotesi che, in altre parole, costituisce un quadro quasi del tutto completo di quegli aspetti dell'universo che lo interessano. La prova d'esame della visione del mondo tanto dello scienziato quanto del mago è la sua capacità di prevedere e, in condizioni opportunamente predisposte, di controllare. Quindi, secondo Bergler, il giocatore coatto vede se stesso come «non la vittima, bensì il braccio esecutivo di un caso imprevedibile» [4].

Ciò che un osservatore esterno considera superstizione del giocatore d'azzardo, è in realtà una manifestazione della ricostruzione che il giocatore, in via d'ipotesi, fa del mondo che il Fato gli ha rivelato, pezzettino per pezzettino. L'esperienza gli ha insegnato, mettiamo, che per vincere egli deve toccare la schiena di un gobbo nel giorno della giocata, tenere nella tasca sinistra una zampa di coniglio, non sedersi al tavolo da gioco con le gambe accavallate e così via. Questo tipo di conoscenza è per lui l'equivalente, diciamo, di quello che la matematica del flusso dell'aria sulle ali può essere per il progettista d'aeroplani.

Poiché le superstizioni del giocatore non hanno in realtà alcuna influenza sul moto dei dadi, sull'ordine delle carte e così via, le sue ipotesi vengono molto spesso falsificate empiricamente. Ogni esperienza falsificante, tuttavia, contiene certi elementi che possono venire integrati nelle linee principali del suo sistema di riferimento ipotetico, salvandone così la struttura complessiva. Il fatto di aver perduto, perciò, non significa che sia sbagliato o inutile, per esempio portarsi dietro una zampa di coniglio; significa soltanto che qualche elemento essenziale per il successo finora è stato trascurato. Magari, l'ultima volta che il giocatore ha vinto c'era una giovane donna bionda dietro alla sua sedia. Aha! Ecco com'è: toccare una gobba, portare una zampa di coniglio, non accavallare le gambe *e in più* avere una giovane donna bionda dietro la sedia. Se anche questo non funziona, il giocatore scoprirà che quella particolare combinazione funziona solo di giovedì, e così via all'infinito. Si aggiungono certi pezzetti di informazione, se ne tolgono altri, e l'intera struttura diventa sempre più complicata. Alla fine, il giocatore davvero controllerà un sistema concettuale che fa concorrenza ai sistemi di conoscenza scientifici, per lo meno in complessità e intricatezza. Sarà diventato un esperto in un mondo assai complicato, aperto soltanto ai pochi iniziati che, con il loro duro lavoro e prendendo dei rischi, ne hanno imparato i misteriosi riti e il linguaggio.

Il mondo magico in cui vive il giocatore coatto non è diverso, in linea di principio, da quelli in cui altri, trascinati anch'essi da fantasie grandiose, cercano di realizzare i loro sogni di potenza. L'astrologia, per esempio, ha costruito un sistema concettuale estremamente complicato, un insieme di teorie e ipotesi che, si sostiene, permettono al «sensitivo» di controllare gli eventi. Il sapere, per esempio, che la congiunzione di certi pianeti in una particolare data porta male per una determinata impresa, mentre qualche altra congiunzione, un altro giorno, porterebbe bene, è un tentativo di controllare gli eventi.

Ma anche le ipotesi dell'astrologia vengono costantemente falsificate dagli eventi. Come è dunque possibile che l'astrologia, come altri sistemi magici

[4] *Ivi*, p. 230.

di quel tipo, rimanga una forza che condiziona la mente degli uomini? Esattamente per gli stessi motivi che valgono nel caso delle ipotesi del giocatore coatto. In primo luogo, qualsiasi contraddizione tra l'esperienza e una delle nozioni magiche viene spiegata riferendosi ad altre nozioni magiche. Così, l'intera struttura del sistema magico di credenze è sostenuta dalla sua stessa circolarità. Questo modo di proteggere il sistema dagli assalti della realtà è particolarmente efficace se le obiezioni si incontrano sempre una per volta, perché in tal caso proprio la dimostrazione che un fatto apparentemente anomalo possa essere incorporato nel sistema serve a convalidarlo. Il giocatore può, per esempio, invocare il fatto di non essersi allacciato bene le scarpe, come sapeva che avrebbe dovuto fare, per giustificare la sua «sfortuna» in un giorno particolare. Questo tipo di spiegazione è formalmente equivalente all'ipotesi, fatta dal programmatore coatto, secondo cui un programma non funziona semplicemente a causa di un errore tecnico di programmazione.

Un secondo modo in cui i sistemi concettuali dei giocatori e dei programmatori vengono protetti è l'elaborazione ciclica. Il giocatore che improvvisamente si rende conto che certi suoi trucchi funzionano solo di giovedì, incorporerà semplicemente questa nuova «intuizione» nel suo sistema di superstizioni già esistente, aggiungendo così in pratica un epiciclo alla sua struttura. Il programmatore è libero di trasformare ciascun nuovo problema imbarazzante in un «caso speciale», da trattare con un sottoprogramma costruito *ad hoc* e quindi incorporato nel sistema globale. Tali illimitate elaborazioni epicicliche dei loro sistemi forniscono tanto ai programmatori quanto ai giocatori un'inesauribile riserva di spiegazioni sussidiarie, buone anche per le più serie difficoltà.

Infine, la stabilità concettuale di un sistema magico o di programmazione può venire protetta negando, per usare le parole di Michael Polanyi,

> a qualsiasi concezione rivale il terreno in cui porre radici. Esperienze a sostegno [dalla concezione rivale] possono solo essere portate una alla volta. Ma una nuova concezione [...] che possa prendere il posto [di quella prediletta] potrebbe venir stabilita soltanto con un'intera serie di esempi pertinenti, e tale evidenza non può accumularsi nella mente [di giocatori o programmatori] se ciascun [elemento di prova] viene a sua volta trascurato perché manca la concezione globale che gli darebbe significato [5].

Il giocatore coatto sfida costantemente le leggi della probabilità; si rifiuta di riconoscerne la validità operativa. Perciò, non può permettere loro di diventare il nocciolo di una visione realistica. Un particolare programma può stare crollando a causa di profonde difficoltà strutturali, matematiche o linguistiche su cui esistono teorie adeguate. Ma il programmatore coatto affronta la maggior parte dei problemi che sorgono con sempre nuovi trucchi di programmazione e così, come il giocatore, si rifiuta di permettere a quei problemi

[5] M. Polanyi, *Personal Knowledge* («Conoscenza personale»). Harper Torchbooks, New York 1964, p. 291.

di enucleare nella sua mente una teoria pertinente. I programmatori coatti, notoriamente, non leggono la letteratura relativa al campo in cui ufficialmente stanno lavorando.

Questi tre meccanismi, chiamati da Polanyi circolarità, autoespansione ed enucleazione soppressa, costituiscono i cardini dell'armamentario difensivo del vero adepto di sistemi magici di pensiero, e in particolare del programmatore coatto. La letteratura psichiatrica ci informa che questa patologia è profondamente legata a deliri di onnipotenza. La convinzione di essere dotato di tutti i poteri, però, non può riposare sugli allori; deve essere costantemente verificata da prove. La prova della potenza è il controllo. La prova della potenza assoluta è il controllo sicuro ed assoluto. Quando parliamo del programmatore coatto, parliamo perciò del suo bisogno di controllo e del suo bisogno di certezza.

La passione per la certezza, naturalmente, è anche una delle grandi pietre miliari di scienza, filosofia e religione. E la ricerca del controllo è intrinseca in tutte le tecnologie. Infatti, la ragione per cui siamo tanto interessati al programmatore coatto è che non vediamo alcuna discontinuità tra le sue motivazioni e il suo comportamento patologico e quelli del moderno scienziato e tecnologo in generale. Il programmatore coatto è soltanto il proverbiale scienziato pazzo cui si è dato un teatro, il computer, ove mettere in scena le sue visioni.

Riprendiamo in esame le tre osservazioni di Bergler sui giocatori. Primo, il giocatore è soggettivamente sicuro di vincere. Lo stesso vale per il programmatore coatto; solo che quest'ultimo, avendo creato un proprio mondo su una macchina universale, ha delle certezze in qualche misura basate sulla realtà. Gli scienziati, con alcune eccezioni, condividono la stessa fede: ciò che la scienza non ha fatto, non l'ha *ancora* fatto; alle domande cui non ha risposto, non ha *ancora* risposto. Secondo, il giocatore ha una fede illimitata nella propria intelligenza. Beh?! Terzo, egli sa che la vita stessa non è altro che un gioco. In modo simile, il programmatore coatto è convinto che la vita non sia che un programma che va su un enorme computer, e che perciò ogni aspetto della vita possa, in ultima analisi, essere spiegato in termini programmabili. Molti scienziati (di nuovo, ci sono significative eccezioni) credono anch'essi che ogni aspetto della vita e della natura possa in ultima analisi essere spiegato in termini esclusivamente scientifici. E infatti, come sottolinea correttamente Polanyi, la stabilità delle credenze scientifiche viene difesa con gli stessi mezzi usati per difendere i sistemi di credenze magiche:

> Qualsiasi contraddizione tra una particolare nozione scientifica e i fatti dell'esperienza verrà spiegata con altre nozioni scientifiche; c'è pronta una riserva di possibili ipotesi scientifiche; capaci di spiegare qualunque evento ci si possa immaginare. [...] *all'interno della scienza stessa*, la stabilità delle teorie nei confronti dell'esperienza viene mantenuta usando epicicli di riserva che sopprimono le concezioni alternative in embrione [6].

[6] *Ivi*, p. 292.

Possiamo quindi riconoscere un *continuum*. Uno dei suoi estremi è occupato da scienziati e tecnologi, che assomigliano molto al programmatore coatto. All'altro estremo stanno quegli scienziati, umanisti, filosofi, artisti e religiosi che cercano la comprensione come persone complete, e da ogni prospettiva possibile. Gli affari del mondo sembrano essere nelle mani di tecnici la cui costituzione psichica è pericolosamente vicina a quella di coloro che occupano il primo estremo. Nel frattempo, sembra che le voci che esprimono la saggezza dell'altro estremo stiano diventando sempre più flebili.

C'è una nota barzelletta che può servire a chiarire questo punto. In una notte scura, un poliziotto si imbatte in un ubriaco. Questi sta inginocchiato, e ovviamente sta cercando qualcosa sotto un lampione. Alla domanda dell'agente, su che cosa stia facendo, risponde che sta cercando le chiavi di casa, che dice di aver perso «laggiù», indicando in direzione dell'oscurità. Il poliziotto gli chiede: «Ma perché, se le ha perse laggiù, le cerca qui sotto il lampione?». L'ubriaco risponde: «Perché c'è molta più luce, qui». È così che anche la scienza procede. È importante rendersene conto, ma è sciocco e inutile prendersela con la scienza per questo. Infatti, quello che si cerca può venir trovato soltanto dove c'è illuminazione. Talvolta, capita perfino di trovare una nuova sorgente di luce nel cerchio illuminato in cui si sta cercando, e di poter così estendere la zona illuminata. Due sono le cose importanti: le dimensioni del cerchio di luce che costituisce l'universo della nostra ricerca, e lo spirito con cui la facciamo. Quest'ultimo deve comprendere un'acuta consapevolezza del fatto che all'esterno c'è oscurità, e che esistono sorgenti di illuminazione di cui si sa ancora molto poco.

La scienza può procedere soltanto semplificando la realtà. Il primo passo in questo processo di semplificazione è l'astrazione. E astrazione significa trascurare tutti quei dati empirici che non si adattano al particolare sistema concettuale all'interno del quale la scienza in quel momento si trova a lavorare; quei dati che, in altre parole, non sono illuminati dalla luce della particolare lampada sotto cui la scienza si trova a cercare le sue chiavi. Aldous Huxley ha commentato con notevole chiarezza:

> Dal punto di vista pragmatico, [gli scienziati] sono giustificati per il fatto di lavorare in questo modo strano ed estremamente arbitrario, perché, concentrandosi esclusivamente sugli aspetti misurabili di quegli elementi di esperienza che sono spiegabili in termini causali, essi sono stati capaci di raggiungere un controllo vasto e sempre crescente sulle energie della natura. Ma potenza non equivale a comprensione e, come rappresentazione della realtà, il quadro scientifico del mondo è inadeguato per la semplice ragione che la scienza non si sogna nemmeno di sostenere di trattare l'esperienza nel suo complesso, ma soltanto alcuni suoi aspetti particolari, in certi particolari contesti. Tutto questo viene compreso ben chiaramente da quegli uomini di scienza che hanno una mentalità più filosofica. Ma purtroppo alcuni scienziati, molti tecnici e la maggior parte dei «*consumatori di gadget*» non hanno avuto il tempo e l'interesse per esaminare i fondamenti e il retroterra filosofico delle scienze. In conseguenza, essi tendono ad accettare la visione del mondo implicita nelle teorie della scienza come un rendi-

conto completo ed esauriente della realtà; tendono a considerare quegli aspetti della realtà che gli scienziati trascurano — perché incompetenti a trattarli — come se fossero in qualche modo meno reali degli aspetti che la scienza ha arbitrariamente scelto come basi per la sua astrazione, tra la totalità infinitamente ricca dei dati di fatto [7].

Una delle più esplicite affermazioni del modo in cui la scienza si propone, deliberatamente e consapevolmente, di distorcere la realtà, per poi procedere ad accettare quella distorsione come «completa ed esauriente» descrizione, è quella fatta dallo scienziato del computer Herbert A. Simon, a proposito del suo stesso orientamento teorico fondamentale:

Una formica, vista come un sistema in azione, è piuttosto semplice. L'evidente complessità del suo comportamento su periodi di tempo abbastanza lunghi è in gran parte un riflesso della complessità dell'ambiente in cui essa si trova. |...| La verità o falsità di |questa| ipotesi dovrebbe essere indipendente dal fatto che le formiche, viste su scala microscopica, siano sistemi semplici o complessi. Al livello cellulare o molecolare, si può dimostrare che le formiche sono complesse; ma questi dettagli microscopici dell'ambiente interno possono essere ampiamente irrilevanti per il comportamento della formica in relazione all'ambiente esterno. È per questo che un automa, benché completamente diverso al livello microscopico, può ciononostante simulare il comportamento complessivo della formica. |...|
Vorrei indagare questa ipotesi, ma con la parola «uomo» al posto di «formica». *Un uomo, visto come un sistema in azione, è piuttosto semplice. L'evidente complessità del suo comportamento su periodi di tempo abbastanza lunghi è in gran parte un riflesso della complessità dell'ambiente in cui egli si trova.* |...| Per parte mia, io credo che l'ipotesi valga anche per l'uomo nella sua interezza [8].

Con un solo colpo di penna, semplicemente sostituendo «uomo» a «formica», la presunta irrilevanza dei dettagli dell'ambiente interno della formica per il suo comportamento è stata elevata a irrilevanza dell'intero ambiente interno dell'uomo per il comportamento umano! Scrivendo 23 anni prima di Simon, ma come se le parole di Simon echeggiassero nelle sue orecchie, Huxley affermava:

A causa del prestigio della scienza come fonte di potenza, e dell'abitudine generale a trascurare la filosofia, la *Weltanschauung* popolare dei nostri tempi contiene una gran componente di quello che si può chiamare *pensiero-nient'altro che*. Gli esseri umani, si suppone più o meno tacitamente, non sono nient'altro che corpi, animali, perfino macchine... i valori non sono nient'altro che illusioni

[7] A. Huxley, *Science, Liberty, and Peace* («Scienza, libertà e pace»), Harper, New York 1946, pp. 35-36.
[8] H.A. Simon, *The Sciences of the Artificial* («Le scienze dell'artificiale»), The MIT Press, Cambridge, Mass. 1969, pp. 24-25.

che in qualche modo si sono venuti a mescolare alle nostre esperienze del mondo; gli accadimenti mentali non sono nient'altro che epifenomeni ... la spiritualità non è altro che... e così via [9].

Salvo che, naturalmente, qui non stiamo parlando della *Weltanschauung* «popolare», bensì di quella di uno dei più prestigiosi scienziati americani. E non è neanche che l'ipotesi di Simon su cosa sia rilevante per l'intero comportamento umano sia «più o meno tacita»; al contrario, egli l'ha espressa (gliene va fatto credito) molto esplicitamente.

Simon ci fornisce anche una descrizione eccezionalmente chiara ed esplicita di come, e quanto interamente, lo scienziato si impedisca di superare la frontiera tra il cerchio di luce proiettato dalle sue stesse supposizioni e l'oscurità esterna. Nel discutere il suo modo di mettere alla prova le tesi che stanno alla base della sua ipotesi, e cioè che l'uomo è piuttosto semplice ecc., egli scrive:

> Ho passato in rassegna alcune delle prove fornite da una grande varietà di prestazioni umane, in particolare da quelle che sono state studiate nel laboratorio di psicologia.
> Il comportamento dei soggetti umani nella soluzione di problemi criptoaritmetici, nel raggiungimento di concetti, nella memorizzazione, nella conservazione delle informazioni in memorie a breve termine, nell'elaborazione degli stimoli visivi, e nello svolgere compiti che usano il linguaggio naturale, sostiene fortemente queste tesi. [...] Generalizzazioni riguardo al pensiero umano... stanno emergendo dall'evidenza sperimentale. Si tratta di cose semplici, proprio come ci aspettiamo in base alle nostre ipotesi. Inoltre, anche se il quadro continuerà a venire allargato e chiarificato, non ci dobbiamo aspettare che diventi significativamente più complesso. Soltanto l'orgoglio umano ci fa ritenere che l'evidente tortuosità del nostro cammino provenga da una fonte del tutto diversa da quella della tortuosità del cammino della formica [10].

L'ipotesi da mettere alla prova qui è, in parte, che l'ambiente interno di un intero uomo sia irrilevante per il suo comportamento. Ci si aspetterebbe che, per metterla alla prova, venisse ricercata dell'evidenza potenzialmente capace di falsificarla. Si potrebbe, per esempio, studiare il comportamento umano davanti al dolore, o a una profonda esperienza religiosa. Ma questi esempi non si prestano facilmente ai metodi di studio dei soggetti umani sviluppati nei laboratori di psicologia. E non è neppure probabile che portino ai risultati semplici che lo sperimentatore, per via delle sue ipotesi, si aspetta. Questi esempi, in realtà, si trovano nella zona oscura in cui il teorico ha perso le sue chiavi; ma sotto il lampione che egli stesso ha eretto c'è un'illuminazione tanto migliore!

Non c'è così alcuna possibilità che l'ipotesi di Simon venga falsificata, nella mente sua o in quella dei suoi colleghi. Il cerchio di luce che determina e delimita il suo raggio visivo semplicemente non illumina alcuna zona in cui

[9] Huxley, *op. cit.*, pp. 36-37.
[10] Simon, *op. cit.*, pp. 52-53.

possano sorgere questioni, mettiamo, di valore o di soggettività. Questioni di quel tipo, essendo — come devono essere — completamente al di fuori del suo universo di discorso, non possono portarlo al di fuori del suo schema concettuale che, come tutti gli altri sistemi esplicativi magici, ha pronta una riserva di ipotesi possibili che spiegherebbero qualsiasi evento immaginabile.

Quasi tutta l'impresa della scienza e tecnologia moderna è afflitta dalla sindrome della ricerca dell'ubriaco, e dalla miope visione che ne è risultato diretto. Ma, come Huxley ha anche indicato, questa miopia non può sostenersi senza essere nutrita da esperienze di successo. La scienza e la tecnologia vengono sostenute dalla loro traduzione in potenza e controllo. I computer e il calcolo, fintantoché possono essere considerati parte di scienza e tecnologia, si nutrono alla stessa tavola. Il fenomeno estremo del programmatore coatto ci insegna che i computer hanno il potere di sostenere visioni megalomaniache. Ma quel potere del computer è soltanto una versione estrema di un potere che è intrinseco a tutti i sistemi di pensiero che si autoconvalidano. Forse stiamo cominciando a capire che i sistemi astratti — i giochi che gli esperti di computer possono generare, nella loro infinita libertà dai vincoli che limitano i sogni dei lavoratori del mondo reale — possono fallire catastroficamente, quando le loro regole vengono applicate sul serio. Dobbiamo anche imparare che lo stesso pericolo è intrinseco in altri sistemi magici che sono ugualmente avulsi dall'autentica esperienza umana, e particolarmente in quelle scienze che insistono nel sostenere di poter catturare *l'uomo nella sua interezza* nei loro disegni schematici astratti.

5
TEORIE E MODELLI

Supponiamo che un gruppo di esploratori di una società altamente tecnologica, proprio come la nostra, ma che non sapesse nulla dei computer, si trovasse di fronte a un computer in funzione. Gli esploratori sono impossibilitati ad aprirlo, non hanno accesso al suo — per così dire — apparato elettroneurofisiologico. Notano, però, che ogni qualvolta essi battono qualcosa sulla sua console a tastiera, le spie del computer lampeggiano in un modo complesso ma palesemente ordinato; talvolta i suoi nastri magnetici ruotano, e la stampante batte un messaggio che sembra essere una risposta a ciò che loro avevano scritto. Dopo un po', si accorgono di poter smontare i nastri magnetici del computer, e di poterne far stampare il contenuto da un'altra macchina, una stampante ad alta velocità, che pure hanno trovato sul luogo dell'esplorazione. Quel contenuto risulta essere leggibile, almeno nel senso che è rappresentato nello stesso alfabeto degli esploratori.

Poiché questa macchina — e gli esploratori riconoscono che è una macchina — è ovviamente uno strumento dotato di comportamento, gli esploratori naturalmente desiderano scoprire le leggi del suo comportamento. Come possono fare per arrivare alla comprensione che desiderano? Di più, cosa può significare il capire le leggi di comportamento della macchina?

Noi, naturalmente, possiamo metterci nella posizione di un osservatore assai privilegiato, più o meno come quella di un professore di chimica che sappia benissimo quali composti ha dato ai suoi studenti da analizzare. Noi sappiamo che la macchina trovata dagli esploratori è un computer, anzi un computer del tal tipo che contiene un programma che pure conosciamo nei dettagli. Siamo perciò in grado di stabilire con precisione il livello di comprensione che gli esploratori avranno raggiunto a ciascuno stadio della loro ricerca. Se, per esempio, essi finissero per produrre un loro computer che, visto dalla nostra prospettiva privilegiata, sembra essere una copia esatta di quello che hanno trovato, e che contiene anche lo stesso programma dell'originale, allora dovremmo ammettere che hanno compreso il computer originale almeno quanto i suoi progettisti.

Tuttavia, anche dei risultati meno eclatanti meriterebbero di esser chiamati comprensione a un livello molto elevato. Supponete, per esempio, che gli esploratori siano riusciti a costruire un computer la cui struttura interna e i cui componenti sono del tutto diversi, mentre il comportamento di input/out-

put è indistinguibile da quello dell'originale; in altre parole, se i due computer non vengono aperti non c'è prova che permetta di determinare quale dei due — quello trovato oppure quello costruito dagli esploratori — abbia generato una certa risposta a un certo input. Può darsi che i componenti interni della macchina trovata siano fatti di fil di ferro, gomma da masticare e nastro adesivo, mentre quelli della copia funzionale costruita dagli esploratori sono tutti elettronici; questo non ha importanza fintantoché, per qualsiasi ragione, la macchina originale non può essere aperta per una dettagliata ispezione interna. (In realtà, naturalmente, può darsi che un simile risultato sia impossibile in linea di principio. Può darsi, per esempio, che il computer originale fosse costruito in modo tale da mettere un punto esclamativo davanti al primo stampato, sempre e soltanto nel diciassettesimo giovedì di ciascun anno bisestile. Anche se questa peculiarità venisse scoperta, gli esploratori non potrebbero mai raggiungere la sicurezza che non vi siano altre stranezze che, benché sistematiche, non sono ancora state scoperte. Noi, come osservatori privilegiati, naturalmente sapremmo tutte queste cose).

Si potrebbe parlare di un buon risultato — benché ancor meno eclatante — anche se gli esploratori riuscissero a costruire una qualche sorta di computer digitale, diciamo una semplice macchina di Turing universale del tipo discusso nel capitolo 2, e se spiegassero quindi la macchina trovata in termini di principi per le macchine di Turing. Saprebbero allora giustificare l'estrema versatilità della macchina trovata, e perfino il fatto che, mettiamo, ci voglia di più per calcolare l'inversa di una matrice grande che di una piccola. D'altra parte, potrebbero essere completamente incapaci di spiegare perché il computer trovato impieghi più tempo per eseguire algoritmi dati in un certo linguaggio di programmazione invece che in un altro. Noi, data la nostra onniscienza riguardo ai computer, sappiamo che la differenza in velocità d'esecuzione è dovuta al fatto che il computer traduce i programmi scritti nel primo linguaggio *una riga per volta* nel suo linguaggio-macchina, per poi eseguire le istruzioni in linguaggio-macchina così generate, mentre con il secondo linguaggio prima traduce l'intero programma in linguaggio-macchina, e poi esegue l'intero insieme di istruzioni così generate. Il secondo processo richiede sempre molto meno tempo del primo. Presumibilmente, gli esploratori finirebbero per elaborare una qualche spiegazione per questo e altri fenomeni enigmatici. Potrebbero, per esempio, ipotizzare che certi linguaggi di programmazione abbiano maggiore familiarità di altri con la macchina, e potrebbero perfino sviluppare una qualche tassonomia dei linguaggi di programmazione basata sulla familiarità della macchina con loro scoperta empiricamente. Il concetto di familiarità, insieme alla tassonomia dei linguaggi per cui è principio organizzatore, diventerebbe allora parte della loro scienza dei computer. È, naturalmente, un concetto debole dal punto di vista della capacità di spiegazione, e può addirittura essere fuorviante. Ma in fin dei conti è molto più facile saperlo per noi, osservatori privilegiati, che per gli esploratori, alle prese con il compito di spiegare fenomeni di complessità illimitata.

Spingiamo la nostra immaginazione un passo più in là: supponiamo che gli esploratori abbiano trovato non soltanto un computer, bensì molti compu-

ter di molti tipi diversi, però tutti macchine cosiddette «a indirizzo singolo». Ricordate che una macchina a indirizzo singolo è quella le cui istruzioni intrinseche hanno la forma «codice di operazione; indirizzo del dato su cui operare» (cfr. p. 91-92). Con un po' di fortuna, e se fossero intelligenti, gli esploratori scoprirebbero quello che senz'altro chiamerebbero «universale del linguaggio» relativo alla struttura grammaticale dei linguaggi-macchina che hanno incontrato. E per spiegarlo altrimenti che non come una pura coincidenza (il che, naturalmente, non sarebbe una spiegazione), essi dovrebbero concludere che questa caratteristica universale di tutti i linguaggi osservati deve essere dovuta a una qualche caratteristica universale, o proprietà innata, della macchina stessa. E sarebbero nel giusto, come noi osservatori privilegiati sappiamo; il fatto che il linguaggio-macchina di un computer abbia il formato a indirizzo singolo è una conseguenza diretta del suo progetto. Infatti, se supponiamo che le macchine trovate dagli esploratori siano normali computer e non robot — vale a dire, che non abbiano percettori, come «occhi» televisivi, né effettori, cioè «braccia» o «mani» meccaniche —, allora tutte le scoperte che gli esploratori fanno e tutte le teorie che sviluppano devono essere basate su osservazioni del, per così dire, comportamento verbale delle macchine. A parte fenomeni minori, benché magari possano aiutare, come il lampeggiare delle spie del computer e il moto occasionale dei suoi nastri magnetici, le uniche prove che i computer offrono della loro struttura sono, dopo tutto, linguistiche. Essi accettano sequenze di dati di ingresso linguistici, sotto forma di testi battuti sui loro terminali e rispondono con dati di uscita linguistici scritti allo stesso modo oppure su nastri magnetici.

Nei capitoli 2 e 3 abbiamo parlato a lungo di mosse lecite in giochi astratti, e di costruzioni grammaticali in linguaggi astratti. Il mio obiettivo era quello di costruire l'idea di computer sulla base di quei concetti. Nella nostra visione attuale stiamo, in realtà, guardando l'altra faccia della stessa medaglia. Ora vediamo che, se si cerca di spiegare i computer con la restrizione di non poterli aprire per guardarci dentro, allora ogni spiegazione dev'essere derivata da basi linguistiche.

La posizione di un essere umano che osserva un altro essere umano non è poi così diversa da quella degli esploratori che vogliono capire il computer che hanno trovato. Anche noi abbiamo un accesso estremamente limitato al materiale neurofisiologico che sembra determinare il nostro modo di pensare. Inoltre, non progrediremmo molto nella nostra attuale comprensione del pensiero anche se potessimo sottoporre il cervello, *in vivo,* al tipo di analisi cui possiamo sottoporre un computer in funzione, studiando cioè le connessioni, seguendo la propagazione di impulsi elettrici, e così via. La nostra ignoranza rispetto alle funzioni cerebrali è attualmente così quasi totale che non potremmo neppure cominciare a delineare le strategie di ricerca adeguate. Staremmo di fronte al cervello aperto, con sofisticati strumenti in mano, più o meno come un lavoratore privo di istruzione potrebbe stare di fronte ai circuiti scoperti di un calcolatore: estasiati forse, ma certamente disorientati. Una microanalisi delle funzioni cerebrali, inoltre, non è utile per capire qualcosa del pensiero più di quanto la corrispondente analisi degli impulsi che scorrono in un com-

puter lo sarebbe per capire il programma che quel computer sta eseguendo. Una simile analisi sarebbe semplicemente fatta al livello concettuale sbagliato. Potrebbe aiutare a eseguire esperimenti cruciali, ma soltanto dopo che quegli esperimenti fossero stati progettati sulla base di teorie elaborate a un livello molto più elevato (per esempio, linguistico).

Poiché, in realtà, gli scienziati soffrono dello stesso tipo di limitazione che abbiamo imposto ai nostri esploratori, e non possono comunicare con un osservatore onnisciente che potrebbe, se soltanto lo volesse, rivelare tutti i segreti, non è sorprendente che almeno alcuni tra gli scienziati cerchino di comprendere come funzionino gli esseri umani più o meno nello stesso modo in cui i nostri esploratori avrebbero potuto cercare di capire il computer da loro trovato, vale a dire, progettando computer il cui comportamento di input/output assomigli quanto più possibile a quello degli uomini.

Qui andrebbe ricordato il lavoro dei linguisti — per esempio quello di Noam Chomsky — anche se esso non coinvolge l'uso dei computer. Una visione semplicistica e grossolanamente fuorviante dell'impresa che la scuola di Chomsky si è proposta è che tale impresa consista nel registrare sistematicamente le regole grammaticali di quante più lingue è possibile. Se questo fosse l'unico, o anche solo il principale, scopo della scuola di Chomsky, ci aspetteremmo che essa pubblicasse una serie di libri, tutti indipendenti uno dall'altro, intitolati «*La grammatica di X*», dove *X* sta per uno dei diversi linguaggi umani conosciuti. In realtà, la più profondamente significativa ipotesi di lavoro di Chomsky è che il patrimonio genetico dell'uomo gli conferisce un insieme di capacità altamente specializzate, e gli impone un corrispondente insieme di limitazioni che, prese tutte insieme, determinano il numero e il tipo di gradi di libertà che regolano e limitano tutto lo sviluppo dei linguaggi umani.

Per capire come una «capacità specializzata» possa essere allo stesso tempo una «limitazione corrispondente», basta che pensiamo a una macchina a indirizzo singolo. Il fatto che una tale macchina sappia decodificare un'istruzione in linguaggio-macchina in termini di una componente (mettiamo, gli otto bit più a sinistra) che è il codice di operazione dell'istruzione, e di un'altra componente (i bit restanti) che è la parte di indirizzo, implica immediatamente che nessun altro formato per le istruzioni è ammissibile per quella macchina. Anzi, l'idea stessa di «grammaticalità» di un intero programma per computer — per non parlare di quella di una singola istruzione in linguaggio-macchina — implica che esistono delle sequenze di simboli che, se possono superficialmente sembrare dei programmi, sono incomprensibili e perciò non ammissibili come programmi. Ciò che da un punto di vista appare come una capacità specializzata di una macchina, da un'altra prospettiva va visto come una restrizione.

Come possiamo allora, si chiede Chomsky, arrivare a una comprensione di quelle capacità ereditate geneticamente che chiamiamo «mente»? Egli risponde che, dato il nostro attuale stato di ignoranza pressoché totale sul cervello vivente, la nostra migliore *chance* è quella di inferire le proprietà innate della mente dai principi altamente restrittivi di una «grammatica universale». Il primo compito del linguista è perciò quello di scrivere delle grammatiche, cioè

degli insiemi di regole, per particolari linguaggi; grammatiche capaci di caratterizzare tutte e sole le frasi grammaticamente ammissibili di quei linguaggi, e poi di postulare principi da cui si possano dedurre le proprietà cruciali di tutte quelle grammatiche. Quell'insieme di principi costituirebbe allora una grammatica universale. L'ipotesi di Chomsky, in altre parole, è che le regole di una simile grammatica universale costituirebbero una sorta di descrizione proiettiva di aspetti importanti della mente umana. Egli non crede, naturalmente, che queste regole si conoscano nello stesso modo in cui si conoscono, mettiamo, le regole per fare lunghe divisioni. Le si conosce, invece (per usare la parola di Polanyi), tacitamente, cioè nello stesso modo in cui si sa come stare in equilibrio quando si corre. Sia nel parlare che nel correre, tra l'altro, una volta che lo si sappia fare lo si fa peggio se si tenta di applicare consciamente delle regole esplicite.

Da un importante punto di vista, quindi, Chomsky è uno dei nostri mitici esploratori. Impossibilitato a esaminare l'interno degli oggetti trovati (le menti umane), e non conoscendo i principi ingegneristici che possono esservi applicati (per esempio la neurofisiologia del cervello vivente) egli si propone di derivarne le leggi dalle manifestazioni del loro comportamento linguistico [1].

Per quanto vasta sia la portata dei propositi della ricerca della scuola di Chomsky, essa è modesta se paragonata a quella dei principali scienziati che lavorano nel ramo della scienza dei computer chiamato «intelligenza artificiale» (*IA*). Herbert A. Simon e Allen Newell, per esempio, che guidano insieme uno dei gruppi più produttivi di ricercatori sull'*IA* alla Carnegie-Mellon University di Pittsburgh, già nel 1958 sostenevano che, per usare le loro parole:

> oggi nel mondo ci sono macchine che pensano, imparano e creano. Per di più, la loro capacità di farlo aumenterà rapidamente finché — nel futuro visibile — il campo dei problemi che potranno trattare avrà estensione paragonabile a quella del campo cui è stata applicata la mente umana [2].

In questo modo, essi proclamavano che l'obiettivo di ricerca della nuova scienza *IA* era niente di meno che la costruzione di una macchina il cui comportamento linguistico, per dire il meno, fosse equivalente a quello umano. Se l'*IA* dovesse raggiungere questo obiettivo, avrà raggiunto il secondo livello — davvero molto elevato — di comprensione delle funzioni umane, per usare termini analoghi a quelli che abbiamo introdotto parlando degli esploratori. In quel contesto, immaginavamo che i nostri esploratori fossero riusciti a co-

[1] Per una discussione breve e di facile comprensione della posizione di Chomsky, vedere il suo *Problems of Knowledge and Freedom* («Problemi di conoscenza e libertà»), Pantheon Books, New York 1971, e in particolare il capitolo I. Discussioni più complete e notevolmente più tecniche si trovano nel suo *Aspects of the Structure of Syntax* («Aspetti della struttura della sintassi»), The MIT Press, Cambridge, Mass. 1965, e nella bibliografia ivi citata.

[2] H.A. Simon - A. Newell, *Heuristic Problem Solving: The Next Advance in Operations Research,* («La risoluzione euristica dei problemi: il prossimo passo nella ricerca sulle operazioni»), *Operations Research*, vol. 6, 1958, p. 8.

struire una macchina il cui comportamento di input/output fosse, di fronte a qualsiasi test, indistinguibile da quello della macchina da loro trovata, benché i componenti delle due potessero essere diversi.

In realtà, gli scopi della ricerca sull'*IA* sono molto più ambiziosi di quelli dei nostri esploratori, che si prefiggevano soltanto di capire come facesse la macchina scoperta a generare i testi di risposta ai testi d'ingresso dati, mentre l'obiettivo dell'*IA* è comprendere come un organismo tratti «un campo di problemi [...] di estensione paragonabile a quella del campo cui è stata applicata la mente umana». Dal momento che la mente umana è stata applicata, per esempio, a problemi di estetica che coinvolgono tatto, gusto, vista e udito, l'*IA* dovrà costruire macchine che sappiano toccare, assaporare, vedere e udire. Dal momento che il futuro in cui il pensiero delle macchine spazierà in modo tanto vasto quanto sostengono Simon e Newell è, al momento in cui scrivo, soltanto «visibile» ma non ancora arrivato, è forse troppo presto per mettersi a speculare su che tipo di equipaggiamento una macchina dovrà avere per poter pensare a preoccupazioni umane come, mettiamo, la delusione negli amori adolescenziali. Ma oggi esistono macchine, principalmente al MIT, alla Stanford University e allo Stanford Research Institute, che hanno braccia e mani i cui movimenti sono coordinati da occhi televisivi computerizzati. Le loro mani hanno dita dotate di cuscinetti sensibili alla pressione, per dar loro un senso del tatto. E ci sono centinaia di macchine che fanno analisi chimiche di routine (e anche non), e di cui si può perciò dire che abbiano il senso del gusto. Macchine capaci di conversare in lingua umana, con una qualità di discorso anche piuttosto elevata, sono state costruite principalmente al MIT e ai Bell Telephone Laboratories. Il Dipartimento della difesa degli Stati Uniti e la National Science Foundation attualmente sostengono sforzi considerevoli verso la realizzazione di macchine capaci di comprendere la voce umana. Chiaramente, l'ambizione di Simon e Newell viene presa sul serio, sia da potenti agenzie governative che da un settore significativo della comunità scientifica.

Posto che non tutti hanno una vista ugualmente buona, non ci si può attendere che tutti, anche oggi, riescano a vedere lo stesso futuro che era già visibile a Simon e Newell nel 1958. E non è neanche necessario che gli psicologi riconoscano la potenza dei modelli della mente umana fatti al computer per poter condividere la grandiosa visione di Simon e Newell. Ci sono segnali molto più modesti a indicare la strada, in modo perfino più diretto.

Tra le altre cose che l'uomo è (e sono molte), egli è anche un organismo dotato di comportamento. Se la sua comprensione di se stesso vuole essere, almeno in parte, scientifica, allora bisogna permettere alla scienza di supporre che almeno alcuni aspetti del comportamento umano obbediscano a leggi che la scienza può scoprire e formalizzare all'interno di qualche schema concettuale scientifico. Per quanto ingenua e informale o, al contrario, sofisticata e formale sia la nozione di «informazione» che si ha in mente, bisogna ammettere che l'uomo agisce su (vale a dire, reagisce a) informazioni che lo raggiungono provenendo dal suo ambiente, e che le sue azioni, in particolare il suo comportamento verbale, a loro volta informano il suo ambiente. Tra le altre cose che l'uomo è, dunque, e di nuovo ce ne sono molte, egli è anche un ricevi-

tore e un trasmettitore di informazioni. Ma anche così, egli è certamente più di un semplice specchio che riflette più o meno fedelmente tutti i segnali che lo investono; infatti, egli presta attenzione soltanto a una piccola frazione di quella che William James chiamava «florida, ronzante confusione» di sensazioni con cui l'ambiente lo bombarda, e trasforma quel distillato del suo mondo in ricordi, immagini mentali di molti tipi, parole e scritti, colpi sulla tastiera del pianoforte; in breve, in pensiero e comportamento. Tra le altre cose che l'uomo è, quindi, e sono molte, egli è anche un elaboratore di informazioni.

In quanto segue, cercherò di sostenere la posizione secondo cui non c'è niente di male nel vedere l'uomo come un elaboratore di informazione (o anche come qualsiasi altra cosa), né nel tentare di comprenderlo partendo da quella prospettiva; a patto, però, di non agire mai come se una singola prospettiva potesse comprendere l'uomo nella sua interezza. Il vedere l'uomo come un sistema per elaborare informazione, di per sé, non lo de-umanizza, e può anche contribuire alla sua umanità, portandolo a una comprensione più profonda di un particolare aspetto della sua natura umana. Potrebbe, per esempio, essere enormemente importante per l'uomo, nel capire la sua spiritualità, conoscere i limiti del potere esplicativo di una teoria dell'uomo come elaboratore di informazioni. Per permetterci di conoscere quei limiti, la teoria dovrebbe, naturalmente, venire svolta in modo assai dettagliato.

Prima di discutere che aspetto potrebbe avere una teoria dell'uomo basata sulla nozione di elaborazione di informazioni, devo dire qualcosa in più sulle teorie, e in particolare sul loro rapporto con i modelli. Una teoria è anzitutto un testo, quindi una concatenazione di simboli di un alfabeto. Ma è anche una costruzione simbolica in un senso più profondo; gli stessi termini usati da una teoria sono simboli che, per parafrasare Abraham Kaplan, cercano a tentoni la loro denotazione nel mondo reale, oppure cessano di essere simbolici [3]. L'espressione «cercare a tentoni» è di Kaplan, ed è una scelta felice, poiché dire che i simboli «trovano» la loro denotazione nel mondo reale negherebbe, o almeno oscurerebbe, il fatto che i termini simbolici di una teoria non possono mai essere fondati in modo conclusivo sulla realtà.

Le glosse che definiscono certe parole in termini di altre parole lasciano queste ultime indefinite. Nella scienza in generale, i simboli vengono spesso definiti in termini di operazioni. Nella fisica, per esempio, la massa è, semplicemente parlando, quella proprietà di un oggetto che determina il suo moto durante le collisioni con altri oggetti (se due oggetti che si scontrano frontalmente mentre si muovono alla stessa velocità si fermano entrambi, si dice che hanno la stessa massa). Questa definizione di massa ci permette di progettare esperimenti che coinvolgono alcune operazioni il cui risultato «misura» la massa degli oggetti. La «quantità di moto» è definita come la massa di un oggetto moltiplicata per la sua velocità (mv), l'accelerazione come la rapidità con cui la velocità cambia nel tempo ($a = dv/dt$), e infine la forza come il prodotto

[3] A. Kaplan, *The Conduct of Inquiry* («La conduzione dell'indagine»), Chandler, San Francisco 1964, p. 296. Questo importante libro contiene un'eccellente discussione di modelli e teorie nelle scienze sociali. Vedere in particolare i capitoli 7 e 8.

della massa per l'accelerazione ($f = ma$). In un certo senso, è sbagliato dire che la forza è «definita» dall'equazione $f = ma$. Una definizione più appropriata, data in alcuni libri di testo, è che la forza è qualsiasi influenza capace di produrre un cambiamento nel moto di un corpo[4]. La differenza tra i due sensi di «definizione» cui si è alluso qui illustra il fatto che le cosiddette «definizioni operazionali» dei termini di una teoria forniscono una base per la progettazione di esperimenti e per la scoperta di leggi generali, ma che queste leggi possono poi fungere da definizioni implicite dei termini che vi appaiono. Questi, e altri, aspetti problematici della definizione implicano che tutti i termini teorici, e quindi tutte le teorie, devono sempre essere caratterizzati da una certa apertura. Non è possibile arrivare alla comprensione completa e finale di un termine di una teoria. Addirittura, per parafrasare ancora una volta Kaplan, può non essere possibile fissare il contenuto di un singolo concetto o termine di una teoria sufficientemente ricca (mettiamo, una teoria della cognizione umana) senza affermare la verità dell'intera teoria.[5] Questo fatto è della massima importanza per qualsiasi giudizio su modelli di fenomeni complessi che usino il computer.

Una teoria, naturalmente, non è soltanto un qualsiasi testo grammaticalmente corretto, che usi un insieme di termini legati in qualche modo simbolicamente alla realtà. È un'aggregazione sistematica di affermazioni di legge. Il suo contenuto, il suo stesso valore come teoria, sta nella struttura delle interconnessioni che collegano le sue leggi tra loro, almeno tanto quanto nelle leggi stesse. (Gli studenti talvolta si preparano per gli esami di fisica mandando a memoria elenchi di equazioni. Può anche darsi che passino gli esami con l'aiuto di simili prodezze mnemoniche, ma non si può dire che conoscano la fisica, cioè che comandino una teoria). Una teoria, o almeno una teoria buona, non è quindi solamente una sorta di banca dati in cui si può «andare a guardare» cosa succederebbe in tali e tal'altre condizioni. È piuttosto come una carta geografica (un'analogia usata anche da Kaplan) di un territorio parzialmente esplorato. La sua funzione è spesso euristica, quella cioè di guidare l'esploratore a ulteriori scoperte. Il modo in cui le teorie rendono diverso il mondo è quindi non tanto rispondendo a domande, bensì guidando e stimolando la ricerca intelligente. E (di nuovo) non esiste un'unica carta «corretta» di un territorio. Una fotografia aerea di una zona svolge una funzione euristica diversa, mettiamo, per un agronomo, rispetto a una mappa demografica della stessa zona. Una utilità delle teorie, dunque, sta nel fatto che esse preparano le categorie concettuali al cui interno il teorico e lo sperimentale faranno domande e progetteranno esperimenti[6].

[4] K.B. Krauskopf - A. Besier, *Fundamentals in Physical Science* («Fondamenti della scienza fisica»), McGraw-Hill, New York 1971, p. 28.

[5] Kaplan, *op. cit.*, p. 57.

[6] Non si deve pensare che questa funzione euristica della teoria sia evidente soltanto nella scienza. Per citare soltanto uno dei possibili esempi al di fuori delle scienze, Steven Marcus, il critico letterario americano, ha usato teorie della critica letteraria, affilate sulla pietra della teoria psicoanalitica, per fare uno studio essenzialmente antropologico di quella sottocultura «straniera,

Di solito, naturalmente, quando parliamo di far funzionare una teoria intendiamo il dedurne delle conseguenze. E con questo, a sua volta, intendiamo il postulare un insieme di circostanze che coinvolga qualche termine della teoria, per poi chiedere cosa la teoria implichi, in quelle particolari circostanze, per gli altri suoi termini. Possiamo, per esempio, descrivere lo stato dell'economia di un particolare Paese a un economista, fornendogli una lista di quegli indicatori economici che compaiono nella sua particolare teoria economica. Forse egli ci farà delle domande che, sostiene, emergono direttamente dalla sua teoria. Queste domande, sia detto per inciso, ci permettono di capire se egli sia, mettiamo, un economista marxista o keynesiano, più delle risposte che ci darà alla fine, in quanto esse rivelano la struttura della sua teoria, la rete di connessioni tra le leggi economiche in cui egli crede. Infine, ci attendiamo di sentire che cosa predice la sua teoria, per esempio che il Paese è in ripresa, o che ci sarà una depressione. Parlando in modo più tecnico, potremmo dire che far lavorare una teoria significa assegnare valori precisi, non necessariamente sempre numerici, ad alcuni dei suoi parametri (cioè, alle entità che i suoi termini rappresentano), e poi stabilire metodicamente quali valori la teoria assegni agli altri suoi parametri. Spesso, ovviamente, arriviamo alle specificazioni su cui vogliamo applicare una teoria interrogando o misurando qualche aspetto del mondo reale. Gli input, per dire così, per una teoria politica, possono per esempio essere stati desunti da un sondaggio d'opinione. Altre volte, le nostre specificazioni possono essere completamente ipotetiche, come per esempio quando chiediamo alla fisica quali sarebbero le conseguenze, sul funzionamento di un orologio, di un lungo viaggio a velocità prossima a quella della luce. In ogni caso, identifichiamo certi termini della teoria con quello che essi significano per noi, associamo loro delle specificazioni e, in pratica, chiediamo alla teoria di calcolare le conseguenze.

Naturalmente, una teoria non può «calcolare» nulla. Dopo tutto, non è che un testo. Ma molto spesso possiamo costruire un modello sulla base di una teoria. E ci sono modelli che possono, in un senso assolutamente non banale, calcolare. Non mi sto riferendo a modelli statici in scala, come quelli che gli architetti fanno per mostrare ai clienti come sarà l'edificio finito. Né penso ai modelli in scala delle ali che gli esperti di aerodinamica provano nei tunnel del vento: anch'essi sono statici. Invece, il sistema formato da un'ala di quel genere e dal tunnel del vento in cui viene provata costituisce un modello del

diversa ed esotica» che fu la sottocultura sessuale dell'Inghilterra vittoriana. Vedere il suo volume *The Other Victorians* («Gli altri vittoriani», Basic Book, New York, 1966). Più recentemente, egli ha scritto nella prefazione a *Engels, Manchester and the Working Class* («Engels, Manchester e la classe lavoratrice», Random House, New York 1974): «Questo mio lavoro può essere considerato come una parte di un esperimento in corso [...] per accertare fino a che punto la critica letteraria possa aiutarci a capire la storia e la società; per vedere fino a che punto la disciplina intellettuale che inizia con una dettagliata analisi testuale possa aiutarci a capire certi documenti sociali, storici o teorici». Nei due libri la teoria della critica letteraria non veniva «applicata» come, per esempio, la teoria chimica può venire applicata all'analisi chimica di un composto; piuttosto, le teorie di Marcus venivano usate euristicamente, come i viaggiatori usano le carte geografiche per esplorare un territorio a loro sconosciuto.

tipo che ho in mente. La sua proprietà cruciale è che esso è capace di comportarsi in modo simile al sistema che vuole rappresentare, vale a dire, a una vera ala d'aeroplano che si muove nell'aria. Il comportamento dell'ala nel tunnel del vento è presumibilmente determinato dalle stesse leggi che governano il comportamento delle ali dei veri aeroplani in volo, lo studioso di aerodinamica, perciò, spera di imparare qualcosa su una vera ala, studiandone il modello in scala ridotta.

Il legame tra modello e teoria sta nel fatto che un modello *soddisfa* una teoria; vale a dire, un modello obbedisce a quelle leggi di comportamento che una teoria corrispondente enuncia in modo esplicito, o che ne possono venire dedotte. Possiamo dire, data una teoria di un sistema *B*, che *A* è un modello di *B* se quella teoria di *B* è anche una teoria di *A*. Accettiamo la condizione, citata anche da Kaplan, che non ci debba essere connessione causale tra il modello e l'oggetto modellato; perché, se si vuole usare un modello come strumento di spiegazione, allora dobbiamo sempre accertarci che ogni lezione che impariamo su un'entità modellata studiandone il modello rimanga anche quando il modello viene rimosso.

La gente, naturalmente, sa dedurre conseguenze dalle teorie senza costruire dei modelli espliciti come, mettiamo, ali in scala ridotta nel tunnel del vento. Ma ciò non significa che quelle conseguenze vengano dedotte senza costruire affatto dei modelli. Quando uno psichiatra applica la teoria psicoanalitica ai dati che il suo paziente gli fornisce, egli sta, per così dire, esercitando un modello mentale — magari molto intuitivo — del suo paziente, un modello costruito in termini psicoanalitici. Per dirla in un modo, l'analista trova utile lo studio del suo modello mentale (*A*) del suo paziente (*B*) per capire il suo paziente (*B*). Per dirla in un altro modo, l'analista crede che la teoria psicoanalitica si applichi al suo paziente, e perciò costruisce un modello di quest'ultimo in termini psicoanalitici, modello cui, naturalmente, si applica anche la teoria psicoanalitica. Egli poi trasforma (è forse meglio dire «traduce») le inferenze ricavate lavorando col modello in inferenze che riguardano il paziente. (Devo aggiungere, per paura di essere altrimenti frainteso, che per quanto lo psicoanalista praticante si affidi alla teoria psicoanalitica e per quanto essa informi il suo atteggiamento, la terapia psicoanalitica consiste solo in piccola parte nell'applicazione diretta o formale della teoria. Nondimeno, è plausibile che tutti noi ricaviamo tutte le nostre deduzioni sulla realtà da modelli mentali la cui struttura — e in larga misura anche il contenuto — è fortemente determinata dalle teorie del mondo cui, in modo esplicito o implicito, crediamo).

I computer rendono possibile un rapporto completamente nuovo tra teorie e modelli. Ho già detto che le teorie sono dei testi. I testi sono scritti in un linguaggio. I linguaggi per computer sono anch'essi dei linguaggi, e possono essere usati per scrivere le teorie. Anzi, per il nostro scopo possiamo evitare di limitare la nostra attenzione ai linguaggi-macchina, o anche ai linguaggi «di livello superiore» di cui abbiamo discusso. Possiamo comprendere tutti i linguaggi, e in particolare anche quelli naturali che i computer siano in grado di interpretare. Il punto è esattamente che i computer *interpretano* i testi che vengono dati loro, in altre parole, che i testi determinano il comportamento dei

computer. Le teorie scritte sotto forma di programmi per computer sono normali teorie, da un certo punto di vista. Un fisico, per esempio, può comunicare la sua teoria del pendolo come un gruppo di equazioni matematiche, oppure come un programma per il computer. In entrambi i casi, egli dovrà identificare i termini della sua teoria — le sue «variabili», in gergo tecnico — con quello cui devono corrispondere nella realtà (potrà dire che *l* è la lunghezza della corda del pendolo, *p* il suo periodo d'oscillazione, *g* l'accelerazione dovuta alla gravità, e così via). Ma il programma per il computer ha il vantaggio non solo di poter essere compreso da chiunque ne abbia studiato il linguaggio, proprio come una formulazione matematica può essere facilmente compresa da un fisico, ma anche di poter essere eseguito da un computer. Se lo si facesse «girare» con un'opportuna assegnazione di valori ai suoi termini, il computer *simulerebbe* un vero pendolo. E da quella simulazione si potrebbero ricavare delle deduzioni, che potrebbero venire tradotte direttamente in deduzioni applicabili ai pendoli reali. Una teoria scritta in forma di programma per computer è così tanto una teoria quanto — se caricata e fatta girare su un computer — un modello cui la teoria si applica. Newell e Simon dicono, a proposito della loro teoria della soluzione di problemi da parte dell'uomo, basata sull'elaborazione di informazioni, che «la teoria svolge i compiti che spiega» [7]. Rigorosamente parlando, una teoria non può «svolgere» nulla. Ma un modello lo può fare, e qui sta il senso della loro affermazione. Dovremo, però, ritornare sulla problematica questione di cosa lo svolgimento di un compito possa e non possa spiegare.

Per aiutarci a intuire che cosa significhi dire che un computer «si comporta», esaminiamo brevemente un modello estremamente semplice: sappiamo dalla fisica — anzi, segue dall'equazione $f = ma$ che abbiamo citato sopra — che la distanza *d* percorsa da un oggetto in caduta libera nel tempo *t* è data da

$$d = at^2/2$$

dove *a* è l'accelerazione dovuta alla gravità. Nei testi di fisica più elementari, si dice semplicemente che *a* è la costante gravitazionale terrestre, vale a dire 9.8 m/sec^2, dove l'unità di distanza è il metro e quella di tempo il secondo. L'equazione stessa costituisce un semplice modello matematico di un oggetto che cade. Se, per semplicità, supponiamo che l'accelerazione *a* sia davvero costante, e cioè 9.8 m/sec^2, possiamo calcolare di quanto un oggetto sarà caduto dopo, mettiamo, 5 secondi: $5 \times 5 = 25$ e $25 \times 9.8 = 245$ e $245:2 = 122.5$. La risposta, come direbbe uno studente di liceo, è 122.5 metri.

I matematici molto tempo fa hanno preso l'abitudine di scrivere le cosiddette «variabili» che compaiono nelle loro equazioni come singole lettere. Forse l'hanno fatto per cautelarsi contro il crampo dello scrittore, o per risparmiare gesso. Qualunque sia la ragione, la loro notazione non è proprio ottimale dal

[7] A. Newell-H.A. Simon, *Human Problem Solving* («Risoluzione umana dei problemi»), Prentice-Hall, Englewood Cliffs, N.J. 1972, p. 10.

punto di vista mnemonico. Dal momento che i programmi per computer sono spesso destinati a essere letti e capiti dalla gente, oltre che a essere eseguiti dai computer, e dal momento che per i computer, entro certi limiti, la lunghezza delle sequenze di simboli che manipolano è indifferente, i programmatori spesso usano delle parole per indicare le variabili che compaiono nei loro programmi. Per altre ragioni, risulta sconveniente usare una giustapposizione di simboli, come in xy, per indicare la moltiplicazione. Molti linguaggi di programmazione usano invece il simbolo «*». Analogamente, «**» viene usato per indicare l'elevazione a potenza. Così, laddove il matematico scrive t^2 il programmatore scrive $t**2$. L'equazione

$$d = at^2/2$$

quando viene trasformata in un'istruzione di programma [8] può così diventare

distanza = (accelerazione *tempo**2)/2

Complichiamo ora appena appena il nostro esempio. Supponiamo che un oggetto venga fatto cadere da una piattaforma statica, per esempio da un elicottero librato a una certa quota sul terreno. L'altezza dell'oggetto sopra il livello del terreno, dopo che è caduto per un certo tempo, sarebbe allora data da

altezza = quota - (accelerazione * tempo**2)/2

Infine, supponiamo che l'elicottero si stia muovendo in avanti a velocità costante, restando sempre alla stessa quota. Se non ci fossero effetti aerodinamici sull'oggetto lasciato cadere dall'elicottero, esso resterebbe esattamente sotto l'elicottero durante tutto il suo viaggio fino a incontrare il terreno. Lo spostamento orizzontale dell'oggetto dal punto in cui è stato lasciato cadere sarebbe perciò uguale allo spostamento orizzontale dell'elicottero da quel punto, e cioè

spostamento = velocità * tempo,

dove con «velocità» indichiamo, naturalmente, la velocità dell'elicottero.

[8] Qui bisogna fare un'importante considerazione tecnica. Benché la «istruzione» scritta qui sia una traslitterazione dell'equazione cui corrisponde, essa non è di per sé un'equazione. In linguaggio tecnico, è una «istruzione d'assegnazione». Assegna un valore alla variabile «distanza». «Distanza», a sua volta, è tecnicamente un «identificatore», il nome di una posizione di memoria in cui viene conservato il valore assegnato alla variabile corrispondente. In matematica, una variabile è un'entità il cui valore non è noto, ma che ha ciò nonostante un valore preciso, che può venire scoperto risolvendo l'equazione. Nei programmi, una variabile può avere valori diversi a diversi stadi dell'esecuzione del programma. Nella matematica ordinaria, per esempio nell'algebra che si studia al liceo, la «equazione» $x = x + 1$ è priva di significato. La stessa sequenza di simboli, se appare come espressione in un programma, ha invece un significato, e cioè che si deve aggiungere 1 al contenuto della posizione di memoria indicata con «x», e sostituire la somma risultante al posto di quel contenuto.

Ora abbiamo da un certo punto di vista due equazioni, da un altro due istruzioni di programma, con cui possiamo calcolare le coordinate, orizzontale e verticale, di un oggetto lasciato cadere da un elicottero in moto uniforme. Possiamo combinarle e inserirle in un piccolo frammento di programma per computer nel modo seguente:

```
FOR tempo = 0 STEP .001 UNTIL altezza = 0 DO;
    altezza = quota - (accelerazione *tempo**2)/2;
    spostamento = velocità * tempo;
    display (altezza, spostamento);
END[9].
```

Questo è un esempio di quella che viene chiamata *istruzione di iterazione*. Dice al computer di fare una certa cosa, finché si arrivi a una determinata condizione. In questo caso, dice al computer di porre anzitutto la variabile «tempo» uguale a zero, poi di calcolare l'altezza e lo spostamento di quello che noi interpretiamo essere l'oggetto in caduta, poi di mostrare le coordinate così calcolate — tra un attimo riparlerò del «mostrare» — e, se l'altezza calcolata non è zero, di aggiungere .001 alla variabile «tempo» e di ripetere il tutto, cioè di iterare il processo. (Questo programma contiene un errore che, per ragioni di semplicità, ho lasciato stare. Così com'è, il programma può girare all'infinito. Per correggerlo, l'espressione «altezza = 0» andrebbe sostituita da «altezza < 0». Lascio al lettore di scoprirne la ragione).

Abbiamo supposto qui che il computer che esegue questo programma abbia in dotazione un apparecchio per esibire i risultati, con le istruzioni corrispondenti. Possiamo immaginare che il *display* del computer sia un tubo a raggi catodici simile a quello di un comune televisore. L'istruzione di display consegna a questo apparecchio due numeri, in questo esempio i valori di altezza e spostamento. Il display fa comparire un puntino luminoso sullo schermo, nella posizione le cui coordinate sono determinate da quei due numeri, cioè tot centimetri sopra e tot centimetri a destra di un fissato punto d'origine.

Se ora facciamo qualche ipotesi aggiuntiva, per esempio sulla persistenza del puntino luminoso sullo schermo e sulla durata complessiva di tutta la storia, possiamo immaginare che il puntino in movimento ci apparirà come un film della caduta dell'oggetto dall'elicottero (cfr. la Figura 5.1). È così possibile, anzi obbligatorio, pensare che il computer «si comporti», e interpretare il suo comportamento come un modello della caduta dell'oggetto.

Sarebbe molto semplice rendere il nostro esempio passo a passo più complicato; dapprima, per esempio, estendendolo fino a coprire la traiettoria di un missile sparato da un cannone e poi, su questa base, estendendolo fino al volo di un satellite orbitante. Avremmo allora descritto almeno le basi più fon-

[9] Abbiamo conservato qui alcune notazioni inglesi del testo originale che sono standard per i linguaggi di programmazione: FOR = per; STEP .001 = il passo .001 verrà usato per le iterazioni temporali del calcolo; UNTIL = finché; DO = itera; display = mostrare (indica qui una routine organizzata in modo tale da produrre un grafico di quelli che sono i suoi «argomenti»); END = fine, (NdT).

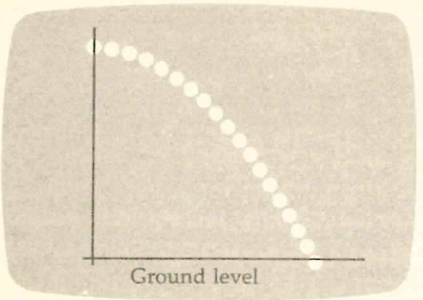


Figura 5.1
Simulazione catodica della traiettoria di un oggetto lasciato cadere da un elicottero in volo.

damentali su cui si fondano le simulazioni orbitali che abbiamo visto spesso in televisione. Ma non è questo il mio scopo. Per quanto semplice sia il nostro esempio, possiamo impararne delle lezioni rilevanti.

Per usare in pratica il modello, uno studioso gli darebbe inizio assegnando dei valori ai parametri quota e velocità, lo farebbe girare sul computer appropriato e osserverebbe il suo comportamento sull'apparato di display del computer. Ci sarebbero, però, delle discrepanze tra le predizioni, per chiamarle così, del modello su un oggetto in caduta e il comportamento della sua controparte reale. Il modello, per esempio, fa implicitamente l'ipotesi che non ci siano effetti aerodinamici sull'oggetto che cade. Ma noi sappiamo che nella situazione reale ci sarebbe sicuramente la resistenza dell'aria. Anzi, se l'oggetto lasciato cadere fosse un paracadute, la vita del suo passeggero dipenderebbe proprio dall'efficacia dell'aria nel rallentarne la caduta. Un modello è sempre una semplificazione, una sorta di idealizzazione di ciò che si intende modellare.

Lo scopo di un modello, ovviamente, è proprio quello di non riprodurre la realtà in tutta la sua complessità. È piuttosto quello di catturare in modo chiaro, e spesso formale, ciò che è essenziale per capire qualche aspetto della sua struttura o comportamento. La parola «essenziale» usata nella frase precedente è di enorme significato, per non dire problematica. Implica, prima di tutto, uno scopo. Nel nostro esempio, vogliamo capire come l'oggetto cada e non, mettiamo, come rifletta la luce del sole durante la sua caduta, o quanto sarebbe profondo il cratere che aprirebbe se venisse lasciato cadere dalla altezza tot. Se ci interessasse quest'ultimo problema, ci dovremmo preoccupare del peso dell'oggetto, della sua velocità all'impatto con il suolo ecc. Noi selezioniamo, per includerle nel nostro modello, quelle caratteristiche della realtà che consideriamo essenziali per il nostro scopo. In situazioni complesse come, mettiamo, modellare la crescita, la decadenza e l'eventuale rigenerazione di una città, l'atto stesso della scelta di ciò che è essenziale e di ciò che non lo è dev'essere, almeno in parte, un atto di giudizio, spesso politico e culturale. E quell'atto deve allora essere basato necessariamente sul modello mentale intuitivo del modellista. Mettendo alla prova un modello, può emergere che qualcosa di essenziale era stato trascurato. Ma, ancora una volta, bisogna esercitare un giudizio per decidere che cosa quel «qualcosa» potrebbe essere, e se sia «essen-

ziale» per lo scopo che il modello si prefigge. Il criterio ultimo, basato com'è su intenzioni e scopi, è in fin dei conti determinato dal modellista individuale, vale a dire umano.

Il problema associato con la domanda su cosa sia e non sia «essenziale» si presta anche ad altre considerazioni. Un modello, dopo tutto, è un oggetto diverso da ciò che si vuole modellare. Ha quindi delle proprietà non condivise dalla sua controparte reale. Gli esploratori di cui parlavamo più sopra possono aver costruito un modello funzionale del computer che hanno trovato usando fibre ottiche e celle fotoelettriche, mentre il computer originale conteneva fili e cancelli elettronici del tipo descritto nel capitolo 3. Essi potrebbero dunque facilmente giungere alla conclusione che la luce sia essenziale per far funzionare un computer. La loro scienza dei computer potrebbe comprendere grandi parti dell'ottica fisica, e così via. È infatti possibile costruire computer usando fibre ottiche ecc. I loro diagrammi logici, cioè il loro progetto scritto, sarebbero fino a un certo punto indistinguibili da quelli dei corrispondenti computer elettronici, poiché gli uni avrebbero la stessa struttura degli altri. Ciò che è essenziale per un computer è l'organizzazione dei suoi componenti e non (di nuovo, fino a un certo punto) il modo preciso in cui quei componenti sono fatti. Un altro esempio: ci sono persone che credono sia possibile costruire un modello computerizzato del cervello umano al livello neurologico. Un modello simile, naturalmente, potrebbe in linea di principio essere descrivibile in termini strettamente matematici. Questo potrebbe far pensare a qualcuno che il linguaggio usato dal nostro sistema nervoso debba essere quello della nostra matematica. Una tale credenza sarebbe un errore del tipo che intendiamo mettere in evidenza qui. John Von Neumann, il grande pioniere dei computer, si soffermò egli stesso brevemente su questo punto:

> Quando parliamo in linguaggio matematico, può darsi che stiamo discutendo di un linguaggio *secondario*, costruito sopra il linguaggio *primario* realmente usato dal sistema nervoso centrale. Così, le forme esteriori della *nostra* matematica non sono rilevanti in assoluto dal punto di vista della valutazione di quale sia il linguaggio matematico o logico usato *veramente* dal sistema nervoso centrale [10].

Una delle funzioni di un modello è quella di mettere alla prova delle teorie ai loro limiti estremi. Ho già detto che il computer può generare film che modellano il comportamento di una particella ai limiti estremi delle velocità relativistiche. Il nostro semplice modello di un oggetto in caduta potrebbe venir usato nella sua forma attuale per simulare, e quindi calcolare, la caduta di un oggetto da una nave spaziale in volo vicino alla superficie della Luna. Tutto quello che dovremmo fare sarebbe riaggiustare l'accelerazione, dandole il valore numerico corrispondente alla gravità che c'è sulla superficie della Lu-

[10] J. von Neumann, *The Computer and the Brain* («Il computer e il cervello»), Yale University Press, New Haven Conn. 1958, p. 82.

na (a patto, naturalmente, che la nave non sia a una quota così elevata sulla superficie della Luna da rendere possibili cambiamenti significativi del campo gravitazionale lungo la traiettoria: un'altra ipotesi implicita). Per questo esercizio di simulazione, non avremmo bisogno di alcuna componente nel modello che corrispondesse alla resistenza dell'aria o ad altri effetti aerodinamici: la Luna non ha atmosfera (ricordate che un astronauta ha lasciato cadere simultaneamente un martello e una piuma sulla superficie lunare, ed entrambi hanno raggiunto il terreno allo stesso istante). È un fatto, però, che il campo gravitazionale lunare varia da posto a posto. Si pensa che queste variazioni siano dovute a cosiddetti *masscon*, ovverosia a certe concentrazioni di materia all'interno della Luna, che agiscono più o meno come degli enormi magneti seppelliti in essa in modo irregolare. L'ipotesi dei masscon è stata avanzata per spiegare le irregolarità osservate nelle traiettorie dei satelliti in orbita intorno alla Luna. Si tratta, in realtà, di un'elaborazione del modello che abbiamo discusso per i corpi in caduta. Il modello elaborato è il risultato della sostituzione di una funzione matematica complessa (in altre parole, una subroutine) al posto del singolo termine «accelerazione» nel nostro modello semplice. Lo dico per illustrare il processo, applicato in questo caso in modo appropriato, di elaborazione di un modello per giustificare nuove e impreviste osservazioni. Ma l'elaborazione dei masscon non era l'unica estensione possibile della teoria o del suo modello al computer. Si sarebbe potuto ipotizzare, per esempio, che la Luna sia circondata da un turbolento mantello di etere, le cui onde e i cui vortici causano il comportamento irregolare delle navicelle spaziali. Ci sono dozzine di buone ragioni per rigettare quest'ipotesi, ovviamente, ma un buon programmatore, se gli si dessero molti dati, saprebbe con maggiore o minore facilità elaborare il modello da cui siamo partiti, aggiungendogli «subroutine per descrivere la turbolenza dell'etere» in modo tale che, alla fine, il modello si comporti proprio come la navicella osservata. Un simile modello, naturalmente, non avrebbe più un aspetto semplice. Anzi, la sua stessa complessità, unita alla precisione cui potrebbe portare i suoi calcoli, gli potrebbe dare una certa credibilità.

In precedenza ho detto che il valore di un teoria non sta tanto nell'aggregato delle leggi che afferma, quanto nella struttura che le interconnette. Il guaio del tipo di elaborazione di modello che risulterebbe da una simile ipotesi di «turbolenza dell'etere» è che essa non fa altro che cucire una «spiegazione» in più su una struttura già esistente. È un rattoppo che non ha radice in nulla di già presente nella struttura. I modelli per il computer, come già abbiamo visto, hanno alcuni vantaggi rispetto alle teorie espresse in linguaggio naturale. Ma queste ultime hanno il vantaggio che i rattoppi sono più difficili da nascondere. Se una teoria scritta in linguaggio naturale è, in realtà, un insieme di toppe su toppe, la sua carenza di struttura sarà evidente dalla sua stessa composizione. Benché un programma per computer costruito in modo simile possa rivelare la sua struttura miserevole a un lettore allenato, questo tipo di difetto non è altrettanto evidente dal funzionamento del programma. Il buon funzionamento di un programma, quindi, non costituisce di per sé una prova adeguata della sua bontà come teoria.

Ho già accennato alla funzione euristica delle teorie. Poiché i modelli in forma di programmi per computer sono anch'essi teorie (almeno, certi programmi meritano di essere considerati tali), ciò che ho detto sulle teorie in generale vale anche per loro, e magari anche con più forza, in questo senso: perché noi possiamo ricavarne delle conseguenze, o anche solo determinarne la coerenza e consistenza, le teorie discorsive devono, come ho detto, essere modellate comunque, cioè essere modellate nella mente. L'eloquenza retorica delle loro affermazioni, particolarmente agli occhi dei loro autori, può dar loro un potere persuasivo che non meritano. Inoltre, può passare molto tempo tra la formulazione di una teoria e il suo venir messa alla prova nella mente umana. I programmi per computer tendono a rivelare i propri errori, e in particolare la loro mancanza di coerenza, in modo rapido ed eclatante. E, in mani capaci, la modellistica al computer fornisce una rapida retroazione (*feedback*) che può avere un effetto davvero terapeutico, proprio grazie alla sua immediatezza. I modelli al computer sono quindi un po' come le istantanee *Polaroid:* è difficile restar convinti di aver fatto una grande foto, quando si ha l'evidenza contraria in mano. Come ha osservato Patrick Suppes:

> Il tentativo di caratterizzare in modo esatto i modelli di una teoria empirica porta quasi inevitabilmente a una comprensione più precisa e chiara dell'esatto carattere di quella teoria. La vacuità e superficialità di molte teorie classiche nelle scienze sociali emerge chiaramente quando si tenta di formulare in modo esatto che cosa costituisca un modello di quelle teorie. Quel tipo di teoria che consiste per lo più in osservazioni ricche d'intuito e in slogan euristici non si presterà a tale trattamento. Lo sforzo di renderla esatta rivelerà allo stesso tempo la debolezza della teoria [11].

Il problema, naturalmente, è proprio sapere quali tipi di teorie si «prestino a tale trattamento».

[11] P. Suppes, *Meaning and Uses of Models* («Significato e usi dei modelli») in B.H. Kazemier and D. Vuysje (a cura di) *The concept and the Role of the Model in Mathematics and Natural and Social Science* («Il concetto e il ruolo del modello nella matematica e nelle scienze naturali e sociali»), Gordon and Breach, New York 1961, p. 172.

6
I MODELLI AL COMPUTER IN PSICOLOGIA

Talvolta, un'idea molto complicata entra nella coscienza collettiva in una forma così semplificata da essere poco più di una caricatura dell'originale; pure, questo semplice abbozzo dell'idea originale può nondimeno cambiare in modo drammatico la concezione popolare della realtà. Come esempio, considerate la teoria della relatività di Einstein. Probabilmente non si capirà mai come e perché questa teoria matematica altamente astratta abbia attirato l'attenzione del grande pubblico, per non parlare del fatto che, per un certo periodo, essa divenne una mania pubblica, e il suo autore un eroe della cultura popolare. Ma lo stesso pubblico che si aggrappò al mito secondo cui soltanto cinque persone al mondo capivano la teoria, riconoscendo così la propria soggezione di fronte a essa, vide la teoria come fonte di una nuova base per il pluralismo culturale. In fin dei conti, la scienza aveva ora stabilito che tutto è relativo. Un esempio più recente può trovarsi nella risposta del pubblico al lavoro di F. Crick e J.D. Watson, che si sono divisi il premio Nobel per la medicina nel 1962 per i loro studi sulla struttura molecolare del DNA, l'acido nucleico che all'interno delle cellule viventi trasmette il patrimonio ereditario. Anche qui, certi risultati altamente tecnici, riportati in un linguaggio assolutamente non comprensibile dall'uomo della strada, sono stati grossolanamente semplificati e generalizzati nella mentalità pubblica, fino a lasciare l'impressione, oggi popolare, che sia già possibile progettare un essere umano dotato di caratteristiche scelte in anticipo. D'un sol colpo, il pubblico si è creato una visione di un'eugenetica positiva basata non su tecniche primitive e (spero) abominevoli come l'uccisione e la sterilizzazione dei «deficienti», bensì sulla creazione, con mezzi tecnologici, di superuomini. Ciò che questi due esempi hanno in comune è il fatto che entrambi hanno introdotto nuove metafore nel buon senso comune.

Una metafora è, nelle parole di I.A. Richards, «fondamentalmente un prestito tra — e un rapporto di — pensieri, un affare tra contesti» [1]. Spesso il valore euristico di una metafora non sta tanto nel fatto che essa esprima una

[1] I.A. Richards, *The Philosophy of Rhetoric* («La filosofia della retorica»), Oxford University Press, Oxford 1936, p. 93.

nuova idea — il che può essere vero o meno — quanto nel fatto di incoraggiare il trasferimento di intuizioni, ricavate in uno dei suoi contesti, nell'altro suo contesto. La sua funzione, così, assomiglia molto a quella di un modello. Uno studente occidentale che si occupi delle società asiatiche può, per esempio, non imparare nulla direttamente dall'osservazione metaforica che i cinesi all'estero sono gli «ebrei dell'Asia». Ma forse non gli è mai capitato di pensare che la posizione degli ebrei nel mondo occidentale, per esempio come imprenditori, intellettuali e bersaglio di persecuzioni, può fungere da modello per provocare intuizioni e domande pertinenti per capire il ruolo e la funzione sociale, mettiamo, dei cinesi in Indonesia. Anche se il richiamare la sua attenzione su questa possibilità forse non darà allo studente occidentale una nuova idea, potrebbe permettergli di ricavare nuove idee dallo scambio dei due contesti, nessuno dei quali gli è nuovo di per sé, ma che tuttavia non aveva mai collegato fra loro prima.

Né l'idea di un oggetto in moto relativo rispetto a un altro, né quella dell'uomo visto fondamentalmente come oggetto fisico, erano sconosciute al buon senso comune negli anni Venti e Sessanta, rispettivamente. Ciò che colpì l'immaginazione popolare quando, per qualche ragione, la stampa si mise a parlare in continuazione della teoria di Einstein, fu il fatto che la scienza aveva sentenziato che la relatività era un fatto fondamentale e universale. Quindi lo slogan «tutto è relativo» fu trasformato in uno schema concettuale legittimo che poteva, almeno in linea di principio, venire accoppiato a qualunque altro universo di discorso, per esempio come un modello esplicativo che legittimava il pluralismo culturale.

I risultati di Crick e Watson caddero su un terreno già preparato dalla vaga comprensione che l'opinione pubblica aveva dei computer, della loro struttura a circuiti e della teoria dell'informazione (con la sua enfasi sui codici e sulla codificazione) e, naturalmente, dalla sua comprensione un po' più accurata della genetica mendeliana, dell'eredità dei caratteri e così via. Fu quindi facile per l'opinione pubblica vedere la «scissione» del codice genetico come un dipanarsi di un programma per computer,e la scoperta della struttura a doppia elica della molecola del DNA come una spiegazione del circuito fondamentale di un computer. Se si accoppia un simile schema concettuale a uno che vede l'uomo come oggetto fisico, è quasi obbligatorio giungere alla conclusione che l'uomo possa essere progettato su ordinazione.

Non serve lamentare che Einstein non abbia mai inteso la sua teoria come una delle due parti della metafora appena descritta. Dopo tutto, è necessario che i due contesti accoppiati da una metafora siano inizialmente disgiunti, proprio come (l'ho sottolineato in precedenza) un modello non deve avere connessioni causali con ciò che si vuole modellare. Il problema con i due usi metaforici che abbiamo citato è che, in entrambi i casi, le metafore sono estese più del lecito. Einstein intendeva dire che non c'è alcun sistema di riferimento fisso, assoluto nello spazio-tempo, al cui interno si svolgano i destini degli eventi fisici. Quindi ogni descrizione di un evento fisico (e, in quel senso, di qualunque cosa) deve essere relativa a un preciso sistema di riferimento nello spazio-tempo. Saltare da questo a «tutto è relativo» è giocare un po' troppo con le

parole. Il contributo di Einstein è consistito nel dimostrare che, contrariamente a quanto si era creduto fino ad allora, il moto non è assoluto. Quando si deduce dalla teoria di Einstein che, mettiamo, ricchezza e povertà sono relative, perché non è il livello assoluto del reddito del ricco e del povero che conta, ma solo il rapporto fra i due, si eleva — senza averne il diritto — una metafora al rango di deduzione scientifica.

L'esempio tratto dalla biologia molecolare illustra una sovraestensione di una metafora in un altro senso; qui, il nostro grado di conoscenza dell'uomo come organismo biologico viene largamente esagerato. Il risultato è, per dire il meno, una chiusura prematura delle idee. La metafora, in altre parole, suggerisce la credenza che tutto ciò che bisogna conoscere sia già noto.

Il computer è diventato una fonte di metafore veramente potenti, e spesso utili. Curiosamente, proprio come con gli esempi già citati, la calda accoglienza del pubblico per la metafora del computer si basa solo su una comprensione estremamente vaga di un concetto scientifico difficile e complicato (in questo caso, la teoria della calcolabilità e i risultati di Turing e Church sull'universalità di certi schemi di calcolo). Il pubblico capisce vagamente — ma è ciò nondimeno fermamente convinto — che qualsiasi procedura efficace possa essere, in linea di principio, eseguita da un computer. Dal momento che l'uomo, la natura e anche la società eseguono procedure che sono sicuramente «efficaci» in un modo o nell'altro, ne segue che un computer può per lo meno imitare l'uomo, la natura, la società in tutti i loro aspetti procedurali. Quindi tutto (di nuovo questa parola!) è almeno potenzialmente comprensibile in termini di modelli al computer e di metafore. Anzi, sulla base di questa generalizzazione non autorizzata delle parole «efficace» e «procedura», si ridefinisce anche la parola «comprensione». Per coloro che sono del tutto preda della metafora del computer, comprendere *x* significa saper scrivere un programma per computer che realizza *x*. Questo è chiaramente esemplificato dal professor Marvin Minsky, direttore del Laboratorio di intelligenza artificiale del MIT, che scrive:

> (I computer) per scrivere della musica veramente buona o per disegnare figure molto significative richiederanno naturalmente dei modelli *semantici* migliori in quei *settori*. Il fatto che questi modelli non siano disponibili non riflette tanto lo stato dei programmi (per computer) euristici, quanto lo stato tradizionalmente disgraziato della critica analitica nelle arti: una conseguenza culturale del fatto che gli analisti estetici per lo più si indignano, quando qualcuno suggerisce che sarebbe possibile capire ciò che stanno cercando di capire [2].

Chiaramente, ciò che Minsky intende con «capire» la musica o la pittura è ben diverso da ciò che, mettiamo, Mozart o Picasso intendevano con lo stesso termine. Uno degli usi della parola «capire» nel passaggio citato è sostanzialmente un gioco di parole — benché, credo, inconsapevole — fatto sull'al-

[2] M. Minsky (a cura di), *Semantic Information Processing* («Elaborazione semantica dell'informazione»), The MIT Press, Cambridge, Mass. 1968, p. 2. Dall'introduzione di M. Minsky.

tro uso. Proprio l'innocenza del suo uso di quella parola testimonia la tenacia della presa che la metafora ha su di lui [3].

Questa nuova definizione di «comprensione» è ora accettata in modo molto diffuso, non solo esplicitamente nei circoli scientifici, ma anche implicitamente nel buon senso comune. Essa implica, come ha rilevato mestamente lo psicologo George A. Miller,

> che l'unica ragione per cui qualcosa non può venir fatto da una macchina di Turing universale è che noi non lo capiamo. Data questa interpretazione di ciò in cui consiste la «comprensione», ogni tentativo di proporre controesempi diventa semplicemente una confessione di ignoranza; oppure, se uno insiste a sostenere di poter capire qualcosa che non sa descrivere esplicitamente, egli diventa un prototipo di quella classe di persone chiamate mistici [4].

In altre parole, la metafora del computer è diventata un altro lampione sotto la cui luce, e soltanto lì, gli uomini cercano la risposta a domande scottanti.

Nessun ramo delle scienze ha eretto questo lampione in modo più deliberato e con più entusiasmo di quanto abbia fatto la psicologia. George Miller scrive:

> Molti psicologi hanno finito per dare per scontato, negli ultimi anni [...] che l'uomo e il computer siano soltanto due diverse specie di uno stesso genere astratto chiamato «sistemi di elaborazione di informazioni». I concetti che descrivono sistemi astratti di elaborazione delle informazioni devono, per forza di cose, descrivere qualsiasi particolare esempio di sistema di quel tipo [5].

Il restringersi del campo visivo che caratterizza la moderna indagine scientifica, proprio come il restringimento del campo visuale permesso da un microscopio, può essere giustificato — e sopportato — soltanto perché ci permette di vedere cose che altrimenti non vedremmo. La scienza e la tecnologia hanno, dopo tutto, conquiste clamorose di cui vantare credito. Data la profondità con cui la metafora del computer ha penetrato la psicologia, è naturale chiedere se in quel campo essa si sia giustificata producendo qualche risultato tangibile.

L'impatto che il computer, nella sua veste di calcolatore numerico ad alta velocità, ha avuto sulla psicologia, benché sia certamente molto grande non conta come «risultato tangibile». La psicologia ha tentato per molto tempo di diventare «scientifica» imitando la scienza che ha avuto il successo più spettacolare, la fisica. Gli psicologi sembrano aver a lungo frainteso proprio che

[3] Va detto che il professor Minsky non ha adottato la metafora del computer in conseguenza di un ingenuo fraintendimento della teoria della computabilità e delle sue implicazioni. Al contrario, essendo un'autorità meritatamente riconosciuta nel campo della scienza dei computer, egli l'ha adottata in modo ponderato e deliberato.

[4] G.A. Miller, *Language, Learning, and Models of the Mind* («Linguaggio, apprendimento e modelli della mente»). Manoscritto non pubblicato, giugno 1972.

[5] *Ivi.*

cosa fosse a rendere la fisica in qualche modo «più scienza» della psicologia. Come anche la sociologia, la psicologia ha interpretato male la proprietà più superficiale della fisica, cioè la sua apparente preoccupazione di usare numeri e formule matematiche, credendo che fosse questo il nucleo che ne faceva una scienza. Larghi settori della psicologia hanno perciò cercato di diventare il più matematici possibile, di contare, di quantificare, di identificare i loro numeri con variabili (preferibilmente quelle con indici greci), e di manipolare le variabili appena trovate in sistemi di equazioni (preferibilmente differenziali) e in matrici, proprio come fanno i fisici. Si è creduto che il solo fatto di profondere energie in questo programma avrebbe senz'altro fatto raggiungere dei risultati utili. La psicometria, per esempio, è e rimane un'attività rispettabile. E non c'è dubbio che la statistica ha beneficiato enormemente dell'esercizio che se ne è fatto in psicologia, proprio come ai tempi della sua infanzia aveva beneficiato della sua applicazione al gioco d'azzardo. Probabilmente ha ripagato in modo simile entrambi i patrocinatori.

Spesso si dice che il computer è soltanto uno strumento. La funzione della parola «soltanto» in questa frase è quella di stimolare l'inferenza che il computer non può essere molto importante in alcun senso fondamentale, poiché gli strumenti in generale non lo sono. Io ho sostenuto la tesi secondo cui gli strumenti informano di sé la ricostruzione immaginaria che l'uomo fa della realtà, e perciò lo istruiscono sulla sua stessa identità. Eppure, il buon senso popolare che percepisce il computer come uno strumento fondamentalmente banale poggia su un'intuizione accurata: il computer, quando è stato usato come «masticatore di numeri» (vale a dire, puramente come un veloce calcolatore di numeri; ed è usato così specialmente nelle scienze comportamentali), ha spesso — come anche George Miller ha osservato — reso muscolose delle tecniche analitiche che sono più potenti delle idee che permettono di esplorare. «I rigoristi metodologici» scrive Stanislav Andreski «sono come dei cuochi che ci mostrino tutti i loro forni scintillanti, i miscelatori, le friggitrici e chissà cos'altro, senza però preparare mai niente che valga la pena di mangiare» [6]. Il masticatore di numeri ad alta velocità è, nelle mani di molti psicologi, semplicemente il loro più nuovo, scintillante, spettacolare frullatore.

Quando il computer è usato semplicemente come uno strumento numerico in psicologia (o in qualunque altro campo), di solito non crea un effetto di focalizzazione della vista; vale a dire, non è paragonabile al microscopio. È perciò irrealistico attendersi che tale uso scopra dei mondi non visti prima, o renda distinto ciò che prima si vedeva soltanto in modo vago. Il punto di vista dell'uomo come specie del più generale genere «sistemi di elaborazione dell'informazione», concentrando la nostra attenzione su un aspetto dell'uomo, ci spinge a rimuovere tutti gli altri suoi aspetti nel buio, oltre il campo che quel punto di vista illumina. Se questo è il prezzo da pagare, abbiamo il diritto di chiedere quali siano i benefici che ci vengono dati in cambio.

Non ci può essere una risposta definitiva a questa domanda, poiché la portata dell'effetto analogico creativo di una metafora deve, per sua stessa na-

[6] S. Andreski, *Le scienze sociali come stregonerie,* Armando, Roma 1977, p. 114.

tura, essere sempre sorprendente e quindi non prevedibile in anticipo. Possiamo dire fin dal principio che il potere di una metafora nel creare nuove intuizioni dipende in larga misura dalla ricchezza degli schemi concettuali che essa fonde, dalla loro reciproca risonanza potenziale. Quanto quel potenziale si realizzerà dipende, ovviamente, da quanto profondamente coloro che partecipano all'atto metaforico creativo possono controllare entrambi i contesti. Questa è, per esempio, la ragione per cui l'esperto di computer che non conosce altro che i computer (il *Fachidiot*, come si dice in tedesco) non può ricavare alcun nutrimento intellettuale di vasta portata dalla sua esperienza, ed è perciò condannato a rimanere per sempre uno hacker. E questa è anche la ragione per cui la metafora del computer è, nelle parole di George Miller, «massimamente produttiva in campi in cui esiste già un notevole fondamento di teoria basato su ricerche precedenti» [7].

Un settore della psicologia era straordinariamente ben preparato a beneficiare di una fusione con il pensiero, orientato verso l'elaborazione tipica degli scienziati del computer; e cioè, il settore che si interessa dei processi cognitivi sottesi dall'acquisizione e memorizzazione dell'informazione. Un enorme lavoro di laboratorio era stato eseguito, per esempio, sul problema della memorizzazione delle cosiddette «sillabe-nonsense». Un tipo di esperimento che è stato condotto da innumerevoli laboratori di psicologia consiste nel far vedere a un soggetto, poniamo, una dozzina di coppie di sillabe formate da tre lettere, una coppia alla volta, e nel chiedergli, a ogni presentazione della prima sillaba di una coppia (tranne la prima), di dire quale sia la seconda. Le sillabe vengono scelte attentamente in modo da essere intrinsecamente prive di senso. Così, per esempio, *TRA* non è una sillaba-nonsense, mentre *GAF* lo è. I soggetti incontrano le coppie della lista ripetutamente, finché sappiano rispondere esattamente a tutte le presentazioni di una sillaba. Edward S. Feigenbaum ha riferito:

> i fenomeni di apprendimento per ripetizione sono ben studiati, stabili e riproducibili. Per esempio, nel tipico risultato comportamentale di un soggetto si trova che:
> a. È più frequente che non ci sia risposta a uno stimolo, piuttosto che un errore manifesto.
> b. Gli errori manifesti sono in generale attribuibili a una confusione, da parte del soggetto, tra stimoli simili o tra risposte simili.
> c. Le associazioni che vengono date correttamente in una serie di prove vengono talvolta dimenticate in seguito, per poi ricomparire e in seguito scomparire di nuovo. Questo fenomeno è stato chiamato oscillazione.
> d. Se prima viene imparato, secondo il nostro criterio, un elenco x di sillabe o coppie di sillabe, e poi viene analogamente imparato un elenco y, e se infine si mette alla prova la memorizzazione dell'elenco x, si scopre che la capacità del soggetto di dare le corrette risposte x è stata degradata dall'apprendimento intercorso nel frattempo. Tale degradamento viene chiamato inibizione retroattiva. Gli errori manifesti fatti nella prova finale su x sono, generalmen-

[7] Miller, *op. cit.*

te, intrusioni dell'elenco *y*. Il fenomeno scompare rapidamente. Di solito, dopo la prima nuova prova, l'elenco *x* è stato imparato di nuovo secondo il criterio.

e. Se le sillabe-stimolo vengono rese più simili, ci vogliono più prove per imparare [8].

Feigenbaum, che attualmente è professore di scienza dei computer alla Stanford University, ha ipotizzato che questo tipo di lavoro di apprendimento coinvolga il soggetto in un processo attivo e complesso di manipolazione di simboli, che potrebbe essere descritto e compreso in modo ottimale in termini di processi più elementari di manipolazione di simboli, proprio del tipo che può venir programmato per un computer.

Naturalmente, niente sarebbe stato più facile che scrivere un piccolo programma per un calcolatore, per permettere a un programmatore di assegnare al computer elenchi di sillabe-nonsense, che il computer avrebbe poi saputo riprodurre perfettamente dopo la prima «prova». L'obiettivo che Feigenbaum si pose era molto più complicato: produrre, sotto forma di programma per computer, un modello di processo cognitivo il cui comportamento globale approssimasse quello dei soggetti umani impegnati a memorizzare sillabe-nonsense, e le cui dettagliate funzioni interne avrebbero rappresentato una spiegazione teorica delle difficoltà effettivamente osservate negli esperimenti. In più, egli voleva che le sue spiegazioni fossero per lo meno consistenti con altre osservazioni psicologiche, come per esempio il fatto che gli esseri umani hanno tanto memorie a breve termine, in cui possono a quanto pare trattenere alcuni simboli da richiamare istantaneamente, quanto memorie a più lungo termine, in cui è possibile conservare una quantità di informazioni quasi illimitata, ma da cui le singole informazioni possono venir richiamate soltanto a spese di qualche sforzo. Se questo «sforzo» per ricordare viene paragonato allo sforzo computazionale necessario per eseguire una subroutine, magari lunga, diventa ovvio come sia possibile incominciare ad applicare la metafora del computer.

L'idea centrale di Feigenbaum è che il computer debba memorizzare una *descrizione* delle sillabe che gli vengono presentate, anziché le sillabe stesse. La sillaba *DAX*, per esempio, può essere descritta dal fatto che la sua prima lettera ha uno spigolo verticale dominante e contiene un ciclo chiuso, la sua seconda lettera contiene una sbarretta orizzontale, e così via. Quando una sillaba viene presentata per la prima volta al sistema, viene aggiunta in memoria una sua descrizione appena sufficientemente dettagliata da permettere di distinguerla dalle altre già memorizzate. Se tale sillaba è uno «stimolo», cioè se è la prima di una coppia di sillabe, allora con essa viene memorizzata anche un'«imbeccata» costituita da una descrizione minimale della sillaba cui è associata. Poiché tutte queste descrizioni sono così minimali, il sistema fa spesso associazioni sbagliate quando gli viene presentato uno stimolo. Ma poiché la

[8] E.A. Feigenbaum, *The Simulation of Verbal Learning Behaviour* («La simulazione del comportamento di apprendimento verbale»), in E.A. Feigenbaum-J. Feldman, (a cura di) *Computers and Thought* («Computer e pensiero»), McGraw-Hill, New York 1963, p. 299.

risposta corretta viene svelata ogni volta che ciò accade, l'informazione descrittiva allora in gioco può essere migliorata aggiungendovi un'ulteriore descrizione. Alla fine, il sistema impara perfettamente l'elenco. Quando si prova con un altro elenco, le descrizioni a esso associate possono di nuovo venir confuse con quelle della prima lista, e viceversa. Il sistema è dunque capace di esibire un'inibizione retroattiva. E, chiaramente, quanto più gli oggetti da imparare vengono resi simili fra loro, tanto maggiore sarà il numero di prove necessario per raffinare il potere discriminatore di ciascun descrittore. Il sistema così si comporta in modo molto simile all'essere umano, di fronte allo stesso compito.

Il programma di Feigenbaum, benché sia oggi molto vecchio (fu completato nel 1959), resta istruttivo almeno da due punti di vista. In primo luogo, ci offre un esempio relativamente semplice di ciò che si intende per modello di processo cognitivo sotto forma di programma per computer. Il modo in cui organizza la sua memorizzazione dell'informazione intende essere una descrizione funzionale della memoria umana a medio termine. Come tale, esso spiega per esempio come sia possibile che noi possiamo dimenticare completamente qualcosa per un lungo tempo, eppure ricordarcene in determinate circostanze. Non è possibile che quell'elemento che si credeva dimenticato fosse stato semplicemente spazzato via dalla nostra mente, perché se così fosse non avremmo potuto mai più avervi accesso. Nel modello di Feigenbaum, nessuna informazione è mai distrutta. Ma l'informazione può venir nascosta, per così dire, da descrittori che portano ad altre associazioni; così, una memoria può schermarne o mascherarne un'altra. Talvolta una raffinazione dell'immagine che fa da schermo (vale a dire, di un'imbeccata) è, nel modello di Feigenbaum, tutto ciò che serve per scoprire (cioè, per rendere di nuovo accessibile) ciò che prima era mascherato.

Il sistema di Feigenbaum richiede anche che le due sillabe da associare siano simultaneamente disponibili al computer (ossia, presenti nella sua memoria) per un breve periodo. Dopo che una «imbeccata» per la voce-risposta è stata associata alla descrizione della sillaba-stimolo, le due sillabe di per sé possono venire cancellate dalla memoria del computer; in altre parole, «dimenticate». C'è quindi una parte del suo sistema che corrisponde in modo plausibile a quel poco che gli psicologi sanno sulla funzione della memoria umana a breve termine. Nessuno, e tanto meno Feigenbaum, sostiene che il suo modello costituisca «la» spiegazione di quei fenomeni. Ma è una spiegazione, in un campo in cui le spiegazioni sono rare.

Il secondo aspetto per cui il programma di Feigenbaum è istruttivo è il fatto che esso si comporta in modi che non sono stati direttamente e deliberatamente «programmati», per usare l'accezione quotidiana di questa parola. Per esempio, il programma mostra quella che gli psicologi chiamano interferenza; vale a dire, l'acquisizione di nuove associazioni interferisce con la produzione di una vecchia associazione, quando le sillabe coinvolte hanno descrizioni molto simili. Il programma non contiene alcuna subroutine messa lì apposta per simulare l'interferenza. Tale fenomeno emerge come conseguenza dell'intera struttura del programma, e costituì una sorpresa per il suo progettista. Da quel

punto di vista, dunque, il modello ha *predetto* un fenomeno comportamentale, il che ne ha accresciuto enormemente la credibilità. Il programma così ci insegna che il facile slogan più volte sentito, «il computer fa soltanto quello che il programmatore gli ha detto di fare» è in un certo senso sbagliato, e in ogni caso problematico.

Il programma che abbiamo discusso fa parte di una classe di programmi chiamati «programmi di simulazione». Il loro oggetto è la simulazione del modo in cui gli esseri umani svolgono certe attività, ma decisamente non l'esecuzione di quei compiti nel modo più efficiente possibile per un computer. Abbiamo osservato, per esempio, che un computer potrebbe facilmente essere programmato in modo tale da «memorizzare» elenchi di sillabe-nonsense alla prima «prova». Ma questo non ci insegnerebbe nulla su come degli esseri umani potrebbero eseguire quello che, almeno superficialmente, sembra essere lo stesso compito.

Poiché i programmi che riguardano gli aspetti cognitivi del comportamento umano ricadono in modo naturale all'interno del campo dell'intelligenza artificiale, *IA* (di cui riparleremo più avanti), bisogna distinguerli da un'altra classe di programmi *IA*, e cioè quelli che sono completamente destinati a svolgere un compito preciso (*task-oriented* in inglese, NdT).

Coloro che lavorano nell'*IA* hanno la tendenza a pensare di lavorare in uno dei due regimi, spesso chiamati *regime di prestazione* e *regime di simulazione (performance mode* e *simulation mode* in inglese, NdT). Forse, il modo migliore per chiarire la distinzione è usare l'analogia con il volo. Praticamente tutti i primi tentativi di comprendere il volo, o di costruire modelli in grado di volare, erano basati sull'imitazione del volo degli uccelli. È un'ipotesi plausibile che il mito di Icaro, l'eroe greco che volò con ali attaccate al corpo con della cera, e precipitò allorché il calore del sole la sciolse, rifletta l'iniziale fallimento dei tentativi dell'uomo di imitare gli uccelli. Potremmo dire che i primi pensatori e pionieri operavano in regime di simulazione. Già alla metà del XIX secolo, tuttavia, uomini come Henson e Stringfellow, e un po' più tardi Langley, passarono a quello che potremmo chiamare regime di prestazione. Partirono dal presupposto che il loro compito era costruire macchine volanti basate su qualunque principio riuscissero a scoprire. Il loro scopo era la prestazione prima di tutto, e la comprensione, soltanto per quanto avrebbe contribuito alla prestazione.

Un terzo regime di operazione andrebbe forse citato in questo contesto: il *regime teorico*. Ci furono grandi teorici dell'aerodinamica prima che ci fossero veri e propri aviogetti. Lord Rayleigh, per esempio, pubblicò lavori importanti imperniati sulla teoria del volo a cominciare del 1875 circa. Ovviamente, dopo che i fratelli Wright ebbero effettuato i loro storici voli nel 1903, l'interesse per l'aerodinamica aumentò continuamente, e non è più sceso fino ai giorni nostri. Ma laddove lo studioso di aerodinamica si dedica alla teoria in quanto tale, e tende a pensare agli aviogetti come a puri modelli delle sue teorie, il progettista di aviogetti vede nella teoria solamente un'altra fonte di idee, che può forse aiutarlo a ottenere prestazioni migliori dalle sue macchine.

La situazione in *IA* è molto simile a quella appena descritta. L'obiettivo della maggioranza tra quanti lavorano in *IA* è costruire macchine che si comportino in modo intelligente, indipendentemente dal fatto che ciò che producono getti o meno luce sull'intelligenza umana. Essi lavorano in regime di prestazione. Vogliono costruire macchine che parlino come gli esseri umani e capiscano il discorso umano, che possano, con l'aiuto di occhi televisivi e di braccia e mani meccaniche, avvitare dadi su bulloni e mettere insieme aggeggi meccanici anche più complicati, che sappiano analizzare e sintetizzare composti chimici, che sappiano tradurre da un linguaggio naturale a un altro, che sappiano comporre musica e programmi complessi per il computer, e così via. Essi sono, naturalmente, lieti di accettare qualsiasi contributo da parte dei teorici (per esempio, gli psicologi) che possa aiutarli a realizzare i loro desideri. Ma il loro obiettivo, a differenza di quello del teorico che va in cerca della comprensione (o sostiene di farlo), è la prestazione, punto e basta.

Un programma come quello di Feigenbaum chiaramente rifugge dalla prestazione; è progettato per richiedere molti tentativi prima di imparare i suoi elenchi mentre, come ho indicato, se il suo scopo fosse la prestazione potrebbe essere modificato in modo tale da memorizzarli al primo tentativo.

Lo spartiacque tra regime di simulazione e regime di prestazione, come ci si potrebbe attendere, non è assoluto. Spesso, l'unico modo per affrontare il problema di far fare un particolare lavoro a un computer è chiedersi come farebbe una persona a farlo. Si comincia così, sostanzialmente, dal simulare il proprio comportamento osservato introspettivamente.

C'è, naturalmente, differenza tra un programma il cui scopo dichiarato è la prestazione ma che, almeno inizialmente, simula «il modo in cui lo farebbe una persona», e un programma che simula quello che una persona farebbe con lo scopo di imparare qualcosa sulle persone. Ma quando un programma intrapreso sotto quest'ultimo vessillo ha successo, allora anch'esso fa una prestazione. Talvolta i suoi autori, a quel punto, non possono resistere alla tentazione di fare prestazioni e di avanzare pretese teoriche su di esse, contribuendo così a rendere meno chiara la linea di demarcazione tra regime di prestazione e regime di simulazione. Newell, Shaw e Simon, per esempio, hanno scritto un programma in grado di dimostrare alcuni teoremi del calcolo proposizionale, simulando il modo in cui studenti ingenui nella logica lottavano per ottenere quelle dimostrazioni [9]. Newell, Shaw e Simon hanno dichiarato come loro scopo: «vogliamo capire come un matematico, per esempio, possa dimostrare un teorema anche se all'inizio non sa come, o se, ci riuscirà». Dopo aver riferito quanto tempo c'era voluto perché il loro programma dimostrasse un certo numero di teoremi, essi osservavano: «Si possono inventare procedure "automatiche" per produrre dimostrazioni [...] ma risulta che queste richiedono tempi di calcolo dell'ordine delle migliaia di anni per la prova di [un particolare teorema]». È difficile vedere in questa affermazione altro che la prete-

[9] A. Newell - J.C. Shaw - H.A. Simon, *Empirical Explorations of the Logic Theory Machine: a Case Study in Heuristics* («Esplorazioni empiriche della macchina per la teoria logica: lo studio di un caso nell'euristica»), The RAND Corp., marzo 1957, Report P-951.

sa che il loro programma possa dare prestazioni utili, oltre a essere magari istruttivo su come facciano «i matematici a dimostrare un teorema». Guarda caso, entro uno o due anni dalla pubblicazione del loro lavoro il matematico Hao Wang ha pubblicato una «procedura automatica», cioè un programma per il computer, capace di dimostrare tutti i teoremi del calcolo proposizionale. Ha dimostrato il particolare teorema, la cui prova Newell, Shaw e Simon avevano stimato richiedere «tempi di calcolo dell'ordine delle migliaia di anni», in 1/4 di secondo su quello che oggi sarebbe considerato un computer molto primitivo [10].

La cattiva definizione della linea che separa il regime di prestazione da quello di simulazione è, con buona giustificazione, un problema che preoccupa poco quanti lavorano nell'*IA*. All'inizio di un grande sforzo di ricerca, ciò che importa è avere un'idea abbastanza chiara almeno del campo generale al cui interno porre le domande o, in altre parole, di quali tra le cose ancora non capite la ricerca si propone di aiutarci a capire. La ricerca di Wang ha prodotto un risultato che ha approfondito la nostra comprensione di certi aspetti della logica matematica. Lo scopo del lavoro riferito da Newell, Shaw e Simon era, nelle loro stesse parole, «comprendere i processi complessi (l'euristica) che sono efficaci nella soluzione dei problemi» [11]. Essi scelsero di esaminare il modo in cui gli uomini dimostrano i teoremi, semplicemente come un esempio di soluzione umana dei problemi. Newell e Simon hanno, come vedremo, continuato fino a oggi il loro lavoro sulla risoluzione dei problemi. Ciò che è stato, e rimane, problematico in questo lavoro è stabilire in che senso della parola «capire» esso ci aiuti a capire l'uomo come elaboratore di informazioni, o come qualsiasi altra cosa. È possibile e utile porsi la stessa domanda riguardo all'intelligenza artificiale in generale.

Il fatto che finora io abbia citato soltanto dei progetti di *IA* molto vecchi non è in alcun senso «scorretto», poiché gli stessi ricercatori sull'*IA* continuano a citare proprio questi esempi (cioè il programma di Feigenbaum e la macchina per la teoria logica di Newell, Shaw e Simon) come lavori fondamentali, mentre procedono nel loro lavoro. Newell e Simon, come ho detto, hanno continuato il loro lavoro sulla risoluzione dei problemi. Nel frattempo altri ricercatori, in particolare quelli del MIT e dei Laboratori di intelligenza artificiale della Stanford University, hanno sempre più rivolto la loro attenzione alla robotica, cioè ai problemi associati con la costruzione di macchine capaci di percepire certi aspetti del loro ambiente, per esempio con l'aiuto di occhi televisivi, e capaci di agire su di esso fisicamente, per esempio per mezzo di braccia e mani meccaniche controllate dal computer. Il loro lavoro, come ci si poteva aspettare, ha generato tutta una serie di sottoproblemi in campi come la visione, la comprensione del linguaggio naturale da parte del computer, il riconoscimento di forme.

[10] H. Wang, *Towards Mechanical Mathematics* («Verso una matematica meccanica») in K.M. Sayre-J.F. Cooson (a cura di), *The Modeling of Mind* («La costruzione di modelli della mente»), University of Notre Dame Press, Notre Dame, Ind. 1963, pp. 92-120.

[11] Newell et al., *op. cit.*

Ovviamente, alcuni di questi sottoproblemi sono anche dei problemi autonomi, vale a dire indipendenti dagli obietttivi di ricerca della robotica. La comprensione del linguaggio naturale da parte del computer è un esempio di problema intrinsecamente interessante, e difficile già di per sé. Il fatto che un progresso in questo campo possa risultare utile per istruire i robot è, per molti dei ricercatori, solamente una motivazione in più, certo non la principale. Riparlerò più avanti della manipolazione del linguaggio naturale da parte dei computer. Per il momento, invece, dedichiamoci all'esame di parte del lavoro più recente sulla risoluzione dei problemi, in particolare quello di Newell e Simon.

La letteratura moderna sulla soluzione dei problemi è scandita da due libri importanti, *How to Solve It* («Come risolverlo») di George Polya e *Human Problem Solving* («Risoluzione umana dei problemi») di Newell e Simon [12]. Il libro di Polya venne pubblicato per la prima volta nel 1945, cioè anni prima che i calcolatori elettronici diventassero comuni strumenti di ricerca. Eppure in esso Polya pone le fondamenta e, in un certo senso, preannuncia tutto il lavoro sulla soluzione dei problemi che sarebbe seguito nei trenta anni successivi. Polya si interessa dei metodi euristici per la soluzione di problemi, cioè di quelle regole empiriche che, quando vengono applicate, possono portare a una soluzione del problema in esame, o a un progresso verso la soluzione, ma non la garantiscono. Le regole euristiche quindi non sono algoritmi, procedure efficaci; sono modi plausibili per aggredire particolari problemi. Polya anticipò molto del lavoro successivo degli scienziati dei computer quando scrisse:

> l'euristica moderna si sforza di capire il processo di soluzione dei problemi, in particolare *le operazioni mentali tipicamente utili in questo processo.* |...| L'esperienza nella soluzione dei problemi e l'osservazione di altri mentre risolvono problemi dev'essere la base su cui l'euristica viene costruita. In questo studio, non dovremmo trascurare alcun tipo di problema, e dovremmo identificare caratteristiche comuni al modo di trattare tutti i tipi di problemi; dovremmo mirare alle caratteristiche generali, indipendenti dall'argomento del problema. Lo studio dell'euristica ha fini «pratici»; una migliore comprensione delle operazioni mentali tipicamente utili nella soluzione dei problemi.|...|
> Va sottolineato che tutti i tipi di problema, in particolare i *problemi pratici* e perfino i *rompicapo* rientrano nel campo dell'euristica |sic|. Va anche sottolineato che l'identificazione di *regole di scoperta* infallibili va al di là dell'oggetto di una ricerca seria. L'euristica descrive il comportamento umano di fronte ai problemi. |...| L'euristica punta alla generalità, allo studio di procedure che siano indipendenti dall'argomento, e si applichino a qualsiasi tipo di problema [13].

Non si sarebbe potuto scrivere uno statuto più chiaro per il lavoro di Newell e Simon. Polya, in pratica, prediceva quegli aspetti, di ciò che Newell e

[12] G. Polya, *How to Solve It* («Come risolverlo»), Copyright 1945 by Princeton University Press, 1957 by G. Polya; Princeton Paperback 1971. A. Newell-H.A. Simon, *Human Problem Solving* («Risoluzione umana dei problemi»), Prentice-Hall, Engewood Cliffs, N.J. 1972.

[13] Polya, *op. cit.*, pp. 129-133 (i corsivi sono di Polya).

Simon avrebbero fatto più tardi, che caratterizzano più profondamente la loro concezione: lo sforzo di capire le operazioni mentali, l'enfasi sulla generalità, sull'indipendenza dall'argomento, e sull'utilità di osservare altri mentre risolvono problemi, nonché la sottolineatura del comportamento durante la soluzione di rompicapo. Infine, Polya sottolinea che il suo libro riguarda i metodi, e che la regola euristica più importante è «il fine suggerisce i mezzi».

Quello che Newell e Simon avrebbero più tardi chiamato «metodo mezzi-fini» venne suggerito per la prima volta quando il modo in cui una versione primitiva della loro macchina per la teoria logica dimostrava i teoremi venne messo a confronto con delle registrazioni di sessioni di «pensiero ad alta voce», protagonisti alcuni studenti, non esperti in matematica, impegnati nella stessa impresa. Questi cosiddetti *protocolli* risultarono molto ricchi di suggerimenti per il lavoro successivo. Redigere protocolli, cioè guardare altre persone risolvere problemi, divenne in pratica un marchio d'autenticità della procedura di Newell e Simon.

La nuova psicologia basata sull'elaborazione di informazioni muove dall'idea fondamentale

> che i computer programmati e i solutori umani di problemi sono entrambi specie appartenenti al genere «Sistemi di elaborazione dell'informazione» |...|
> Quando cerchiamo di spiegare il comportamento dei solutori umani (o dei computer), scopriamo che la loro flessibilità — la loro programmabilità — è la chiave per comprenderli. La loro vitalità dipende dalla capacità di adattarsi a una grande varietà di ambienti. |...|
> Se con attenzione mettiamo in evidenza le influenze degli ambienti rispetto alle componenti fisiche fondamentali e all'organizzazione, riveliamo la vera semplicità del sistema adattabile. Perché, come abbiamo visto, dobbiamo postulare soltanto un sistema di elaborazione di informazioni molto semplice, per poter spiegare il modo umano di risolvere i problemi in situazioni come il gioco degli scacchi, la logica e la criptoaritmetica. Il comportamento apparentemente complicato del sistema di elaborazione di informazioni in un dato ambiente è prodotto dall'interazione tra le esigenze dell'ambiente e pochi parametri fondamentali del sistema, in particolare le caratteristiche delle sue memorie.
> Le cose sono semplici non perché la legge dei grandi numeri faccia cancellare molti termini, bensì perché i termini si allineano in una catena mezzi-fini in cui solo i punti-fine contano (cioè, l'equifinalità) [14].

Questa è un'affermazione davvero notevole, specialmente alla luce della pretesa di Simon che l'ipotesi che rappresenta «valga anche per l'uomo nella sua interezza». Vale la pena che tentiamo di capire cosa sia questo sistema di elaborazione di informazioni «molto semplice» che produce un comportamento complesso in funzione del suo ambiente e di «pochi parametri fondamentali». Ci dobbiamo anche chiedere in che modo problemi come gli scacchi, la logica, la criptoaritmetica possano venire generalizzati all'«intero campo |di problemi|

[14] Newell-Simon, *op. cit.*, pp. 870-871.

cui è stata applicata la mente umana», vale a dire a quel campo cui questi stessi autori hanno promesso che il computer verrà applicato «nel futuro visibile» [15]. Quest'ultima domanda è particolarmente pertinente, poiché i programmi euristici esistenti per la soluzione dei problemi trattano soltanto problemi molto semplici di scacchi e di logica. La criptoaritmetica qui non conta, poichè è quello che viene chiamato, perfino nei circoli dell'*IA* «problema-giocattolo»[16].

Abbiamo già stabilito che è del tutto appropriato, e anche utile, partire da un punto di vista molto particolare e, da quella prospettiva, vedere l'uomo come elaboratore di informazioni. E dal momento che il computer, la macchina di Turing, è un elaboratore di informazioni universale, è naturale paragonare l'uomo, visto da quella prospettiva, al computer. La *psicologia* dell'elaborazione di informazioni, però, non ne è una *neurofisiologia*. Non tenta spiegazioni in termini di bit, o facendo analogie con flip-flop, circuiti elettronici ecc. Evita, e a ragione, perfino spiegazioni che dipendano da un paragone con le manipolazioni di simboli tipiche delle macchine di Turing classiche, per esempio scrittura, lettura, cancellazione e confronto di simboli estremamente semplici e irriducibili, come «0» e «1».

Ricordate che un programma per un particolare computer è sostanzialmente una descrizione di un altro computer, che trasforma la prima macchina nella seconda. Si può perciò progettare un computer, e in seguito realizzarlo sotto forma di programma per computer, i cui processi elementari d'informazione interni operano su strutture simboliche arbitrariamente complesse, cioè leggono, scrivono, cancellano, confrontano fra loro quelle strutture simboliche, e così via. Con strutture di quel tipo si possono rappresentare formule della logica, espressioni matematiche, parole, frasi, disegni architettonici ecc., e naturalmente programmi per computer che possono a loro volta essere manipolati da processi informatici elementari.

Un metodo di programmazione particolarmente utile, per esempio, consiste nell'organizzare l'informazione sotto forma di concatenazioni di elementi individuali. Il «legame» che concatena un elemento al successivo è un indirizzo di macchina, o indicatore (*pointer* in inglese, NdT), che viene memorizzato a fianco dell'elemento in questione e punta verso il successivo. Una lista (come viene chiamata di solito una simile catena) di elementi concatenati in questo modo può a sua volta essere considerata un elemento, e può così venir «puntata» da un altro elemento ancora. In questo modo si possono creare e manipolare strutture di enorme complessità. Certi processi informatici elemen-

[15] H.A. Simon - A. Newell, *Heuristic Problem Solving*, cit., pp. 8 e ss.

[16] È vero che esistono programmi molto potenti per fare della logica simbolica e della matematica estremamente complesse. Ma questi sono programmi specializzati che — benché possano avere approfittato di tecniche di *IA* all'inizio del loro sviluppo — non possono in alcun modo essere paragonati ai sistemi di elaborazione delle informazioni di cui parlano Newell e Simon. È fuori discussione che essi non sono, per la psicologia, più importanti dei molti programmi che risolvono sistemi di equazioni differenziali. Anche i programmi attualmente più potenti per giocare a scacchi sono stati scritti in regime di prestazione e, benché possano usare occasionalmente tecniche dell'armamentario *IA*, sono anch'essi sostanzialmente irrilevanti per la psicologia.

tari possono trattarle come elementi singoli, mentre altri possono scorrere al di sopra di esse, per esempio inserendo o distruggendo delle sottostrutture.

Un sistema di elaborazione dati è perciò, in questo contesto, un sistema di calcolo a componenti fisici (*hardware*) unito a un programma in grado di eseguire processi informatici elementari su strutture simboliche memorizzate. Ha, naturalmente, apparecchi di input/output, come telescriventi o altro, che permettono un'adeguata comunicazione con il mondo al suo esterno.

Il più ambizioso sistema di elaborazione dell'informazione che sia stato costruito allo scopo di studiare il comportamento umano nella risoluzione di problemi è il Risolutore Generale di Problemi (o *GPS*, dall'inglese *General Problem Solver*) di Newell e Simon [17].

> I metodi principali del *GPS* incarnano insieme l'euristica dell'analisi mezzi-fini. |...| L'analisi mezzi-fini è tipizzata dal seguente esempio di ragionamento basato sul buon senso comune:
>
> Voglio portare mio figlio all'asilo. Qual è la differenza tra ciò che ho e ciò che voglio? La distanza. Che cosa fa cambiare la distanza? La mia macchina. La mia macchina non funziona. Di che cosa ha bisogno per funzionare? Di una batteria nuova. Chi ha batterie nuove? L'officina di riparazioni. Voglio che l'officina di riparazioni metta una batteria nuova nella mia macchina. Ma i meccanici dell'officina non sanno che ne ho bisogno. Dove sta la difficoltà? Nella comunicazione. Che cos'è che permette la comunicazione? Il telefono ... e così via.

Questo tipo di analisi — che classifica le cose a seconda della funzione che svolgono, e oscilla tra i fini, le funzioni necessarie, e i mezzi per svolgerle — costituisce il sistema fondamentale dell'euristica del *GPS*.
Più precisamente, questo sistema mezzi-fini presuppone quanto segue:

1. Se si dà un oggetto che non è quello desiderato, sarà possibile riconoscere delle differenze tra oggetto dato e oggetto desiderato.
2. Gli operatori influenzano certe caratteristiche dei loro operandi, lasciandone altre immutate. Quindi, gli operatori possono venire caratterizzati dai cambiamenti che producono, e possono essere usati per cercare di eliminare le differenze tra gli oggetti cui vengono applicati e quelli desiderati.
3. Se un operatore desiderato non è applicabile, può essere vantaggioso modificare i suoi input in modo che lo diventi.
4. Alcune differenze risulteranno essere più difficili da influenzare di altre. È vantaggioso, quindi, tentare di eliminare le differenze «difficili», anche a costo di introdurre nuove differenze di minore difficoltà. Questo processo può venire ripetuto fintantoché si stiano facendo progressi verso l'eliminazione delle differenze più difficili [18].

[17] Questo sistema è descritto in dettaglio in Newell-Simon, *op. cit.*, specialmente nel capitolo 9, *Logic: GPS and Human Behaviour* («La logica: il GPS e il comportamento umano»), pp. 455-554.

[18] *Ivi*, p. 416.

Per vedere come questo metodo funzioni su un problema del tipo di quelli su cui il *GPS* è stato effettivamente applicato, considerate il seguente indovinello criptoaritmetico:

$$\begin{array}{r} DO \\ +\ IT \\ \hline TTD \end{array}$$

A un soggetto viene detto che questa è una codificazione di una normale addizione. Ogni lettera rappresenta una cifra, e lettere diverse non possono rappresentare la stessa cifra. Il suo compito è associare cifre alle lettere, in modo tale che l'espressione data rappresenti un'addizione corretta. Egli deve produrre un protocollo, cioè dire ad alta voce quello che sta pensando. Quello che segue è un possibile esempio di protocollo, frammezzato da un'analisi in termini del tipo *GPS*:

Soggetto: $D+I$ deve essere maggiore di 9, perché nel totale c'è un riporto nella colonna più a sinistra.

Analisi: il soggetto ha applicato l'operazione «elabora la colonna».

Soggetto: T deve essere 1, poiché è un riporto.

Analisi: il soggetto ha applicato l'operatore «assegna un valore». Egli ha raggiunto un obiettivo parziale, e ha ridotto la differenza tra oggetto dato e oggetto desiderato. Il nuovo «oggetto dato» è ora

$$\begin{array}{r} DO \\ +\ I\ 1 \\ \hline 1\ 1D \end{array}$$

Soggetto: O deve essere almeno 2.

Analisi: il soggetto ha applicato l'operatore «genera possibili valori» a O. (Ci deve essere stato un tacito tentativo di applicazione dell'operatore «assegna un valore», il cui risultato è stato respinto).

Soggetto: proviamo con $O=2$.

Analisi: il soggetto ha applicato l'operatore «assegna un valore». È un'altra riduzione di differenze. Il nuovo «oggetto dato» è ora

$$\begin{array}{r} 32 \\ +\ I1 \\ \hline 113 \end{array}$$

Soggetto: $I = 8$.

Analisi: con l'applicazione dell'operatore «assegna un valore», la differenza tra oggetto dato e oggetto desiderato viene eliminata. L'obiettivo è raggiunto.

Questo era un problema molto più semplice di quelli assegnati tipicamente ai soggetti e al *GPS*. Un esempio molto più realistico di problema che è stato completamente analizzato è:

$$\begin{array}{r} DONALD \\ +\ GERALD \\ \hline ROBERT \end{array}$$

ove si sappia che $D = 5$. L'esempio che abbiamo analizzato ha inoltre il demerito di non mostrare alcuna mossa sbagliata, alcun ritorno indietro e così via. Ciononostante esso dà un'idea generale, benché pallida, del modo in cui il *GPS* funziona, e di cosa sia un protocollo.

Va anche compreso che il *GPS* non è il modello della teoria di Newell e Simon. Il *GPS* ha implicazioni maggiori, riguardo a un livello di generalità distinto e indipendente dai compiti da eseguire, rispetto a quanto richiesto dalla loro teoria. Anzi, non esiste alcun programma per computer che sia un modello della loro teoria. Esiste invece un certo numero di programmi — certo non tutti compilati da Newell e Simon, o dai loro collaboratori — che sono sostanzialmente in accordo con la teoria e impiegano i «metodi pricipali del *GPS*» elencati sopra. È la teoria dell'uomo come elaboratore di dati che ci interessa qui, non il *GPS* in quanto tale. E quella teoria ci interessa esattamente perché essa, in una variante o in un'altra, talvolta esplicitamente e talvolta implicitamente, sta dietro a quasi tutta la nuova psicologia dell'elaborazione di informazioni, e costituisce praticamente un dogma per la comunità di scienziati che lavorano sull'intelligenza artificiale.

Le conclusioni fondamentali cui la teoria giunge sono le seguenti:

> Tutti gli esseri umani sono dei sistemi di elaborazione di informazioni, e hanno quindi certe caratteristiche basilari di organizzazione in comune; tutti gli esseri umani hanno in comune alcune caratteristiche strutturali universali, come per esempio parametri di memoria quasi identici. Queste comunanze producono delle caratteristiche di comportamento comuni a tutti gli umani risolutori di problemi.

> Dal momento che il sistema di elaborazione dell'informazione [cioè, l'essere umano visto come sistema di elaborazione dell'informazione, NdA] può essere visto come il risultato di (1) struttura di base e (2) contenuto della memoria a lungo termine (cioè, programmi e dati), ne segue che ogni proposta di comunanza tra solutori di problemi che non sia attribuibile direttamente alla struttura di base deve essere rappresentata come identità o similarità nel contenuto delle memorie a lungo termine, nel sistema di produzione o in altre strutture di memoria immagazzinate.
> (La teoria) propone un sistema che, dato un tempo sufficiente, può assorbire qualsiasi specificazione: può diventare comprensivo di tutti i dettagli, mettiamo, di un'enciclopedia (o di una biblioteca di enciclopedie). Quindi, la teoria affida la determinazione di differenze e somiglianze di comportamento direttamente alle cause che definiscono il contenuto che sarà immagazzinato nella memoria umana a lungo termine. Ma questi determinanti del contenuto sono in larga misura contingenti, dipendono dai dettagli dell'esperienza di vita dell'individuo. Questo non significa che i processi determinanti siano arbitrari, o capricciosi, o sregolati. Significa che il contenuto sarà vario quanto il campo dei fenomeni fisici, biologici e sociali che circondano l'individuo e da cui egli li estrae [19].

Ciò che è notevole in queste conclusioni è la loro portata, per esempio il fatto che il sistema che la teoria propone, presumibilmente un sistema del tipo *GPS*, può assorbire qualsivoglia specificazione.

Questa pretesa — come chiamarla altrimenti? — è coerente con altre pretese degli autori, e cioè che la teoria possa spiegare l'uomo nella sua interezza e che i computer tratteranno, in un futuro già in vista, problemi in tutto il campo del pensiero umano. L'assurdità di quanto si pretende per un sistema del tipo *GPS* è sottolineata dall'affermazione di Newell e Simon secondo cui «Il comportamento apparentemente complicato del sistema di elaborazione dell'informazione in un dato ambiente è prodotto dall'interazione tra le esigenze di quell'ambiente e pochi parametri fondamentali del sistema, in particolare le caratteristiche delle sue memorie» [20]. Questo, naturalmente, è del tutto coerente con la loro convinzione che l'uomo (come la formica) sia «piuttosto semplice». Ma in questo contesto viene avanzata una pretesa tecnica, e cioè che il *GPS* sia piuttosto semplice, nel senso che, cambiando *pochi* dei suoi parametri, la sua interazione con l'ambiente produrrà un comportamento opportunamente vario, e che approssima quello dell'uomo.

Nel normale linguaggio tecnico si dice che un sistema è sensibile a «pochi parametri» quando tutti gli aspetti importanti del suo comportamento posso-

[19] *Ivi*, pp. 864-866. Queste affermazioni provocano a un confronto con quelle di B. F. Skinner: «Un'analisi scientifica del comportamento deve, io credo, presupporre che il comportamento di una persona sia controllato dalla sua storia genetica e ambientale, piuttosto che dalla persona stessa come agente iniziale, creativo» (dal suo *About Behaviorism*, "Sul comportamentismo", New York, A.A. Kopf, 1974, p. 189). L'unica differenza tra la posizione di Skinner e quella della teoria in discussione — differenza importante da un certo punto di vista, ma del tutto irrilevante da un altro — è che Skinner si rifiuta di guardare dentro quella scatola nera che è la persona, mentre la teoria vede questo interno come un computer.

[20] *Ivi*, p. 809.

no essere completamente determinati girando poche manopole, o introducendo pochi dati nelle sue memorie. Un navigatore computerizzato è di questo tipo, per esempio. Esso può guidare una nave dovunque, date solo le coordinate geografiche della destinazione, alcune informazioni meteorologiche e così via. Ma la conversione di un sistema del tipo *GPS* da giocatore di scacchi a, mettiamo, solutore di indovinelli criptoaritmetici non si può attuare cambiando soltanto pochi numeri. In pratica, l'intera «struttura di memoria» del *GPS* deve essere sostituita ogni qual volta esso viene indirizzato a un nuovo compito. In altre parole, il *GPS* in sostanza non è altro che un linguaggio di programmazione in cui è possibile scrivere programmi per certi compiti altamente specializzati. Ma, a meno di dover considerare un programma come parametro singolo, il *GPS* non costituisce affatto un sostegno per la pretesa che la complessità del comportamento umano sia funzione soltanto dell'ambiente umano, e di pochi parametri interni al sistema umano di elaborazione dell'informazione.

Occasionalmente, Newell e Simon usano in effetti una nota di cautela, come per esempio quando ammettono che «non sappiamo quale parte di tutta l'attività umana di soluzione di problemi impieghi uno "spazio dei problemi", ma tra i compiti e tra gli individui che abbiamo studiato — uno spettro abbastanza esteso da far pensare che le comunanze non siano casuali — uno spazio dei problemi viene usato sempre». Ma poi, smentite blande come questa sono compensate da affermazioni come: «Nonostante la base evidenziale esplicita della teoria abbia un'estensione limitata, noi la proporremo come teoria generale della soluzione di problemi, senza cercare di stabilire le frontiere della sua applicabilità» e «noi crediamo che la teoria che stiamo proponendo abbia validità più estesa dei dati particolari su cui l'abbiamo eretta» [21].

È proprio questa pretesa ingiustificata di universalità a degradare il loro uso dei computer, dei sistemi di calcolo, dei programmi ecc., dal livello di teoria scientifica a quello di metafora. Lo dicono essi stessi: «Qualcosa cessa di essere metafora quando se ne fanno calcoli dettagliati; resta metafora, quando è di per sé ricca di tratti distintivi, la cui rilevanza per l'oggetto di confronto è problematica» [22]. La domanda quindi è, si possono fare calcoli dettagliati a partire dalla loro «teoria»? (Continuerò a usare la parola «teoria», perché sarebbe troppo scomodo scrivere sempre «supposta teoria» quando si parla del lavoro in questione).

La risposta, a prima vista, sembra essere un sonante «sì». Il libro di Newell e Simon non è forse pieno di esempi di calcoli fatti col *GPS*? Ma qui c'è una sottigliezza, di grande importanza, quasi sempre trascurata da coloro che lavorano nell'intelligenza artificiale, credendo anch'essi di essere in possesso di vere teorie. Forse il modo più chiaro di illuminare questo punto è confrontare il simulatore della memoria ripetitiva di Feigenbaum con i programmi *GPS* riportati nel libro di Newell e Simon. Il programma di Feigenbaum è, come ho detto in precedenza, un modello di teoria psicologica, cioè di come la gente

[21] *Ivi*, p. 791, 809.
[22] *Ivi*, p. 5.

affronta il problema di memorizzare sillabe prive di senso. Il programma stesso è anche una teoria, come ho messo in evidenza; per esempio, se lo si dà a uno psicologo che ha familiarità con il linguaggio di programmazione in cui è scritto, ci si aspetta che egli lo capisca. La proprietà che lo qualifica come teoria, però, è il fatto di enunciare principi da cui si possono ricavare conseguenze. Questi principi sono essi stessi in forma di programmi per computer, e le loro conseguenze emergono nel comportamento del programma, cioè nella lettura che il computer fa del programma. Tra queste conseguenze ci sono i noti fenomeni di interferenza e inibizione retroattiva che ho citato in precedenza.

La situazione è del tutto diversa quando, mettiamo, si fa girare il programma di teoria logica nel *GPS*. Di sicuro, quel programma è di nuovo una teoria (benché piuttosto banale), e precisamente una teoria di come dei novizi si comportano dovendo risolvere certi problemi elementari di logica. Ma il *GPS*, e questo è il punto cruciale, *è soltanto uno schema al cui interno il programma di teoria logica viene eseguito*. Il *GPS* è, in pratica, un linguaggio di programmazione in cui è particolarmente facile scrivere programmi di teoria logica, programmi di criptoaritmetica, e così via. I processi informatici elementari che ne costituiscono le istruzioni elementari sono semplicemente le istruzioni primitive della macchina in cui il *GPS* ha trasformato il suo computer-ospite. Il *GPS* in quanto tale non contiene alcun principio; a meno che non si considerino principi osservazioni come quella per cui, per risolvere problemi, è necessario operare in termini di strutture simboliche molto generali che rappresentano oggetti, operatori, peculiarità degli e differenze tra gli oggetti, oppure è necessario istituire una biblioteca di metodi, e così via. Anche allora, il *GPS* non permette di ricavare delle conseguenze da tali «principi».

Dire che il *GPS* è, in un qualunque senso, un'incarnazione di una teoria della soluzione umana di problemi, è come dire che l'algebra delle scuole superiori è un'incarnazione di quel tipo. Anche quest'ultima è un linguaggio, uno schema di calcolo, al cui interno si può rappresentare una teoria cui si sia già arrivati per altra strada. Esiste, naturalmente, una teoria dell'algebra. Ed esistono teorie dei linguaggi di programmazione. Ma nessuna di queste pretende di dire alcunché sulla psicologia della risoluzione umana dei problemi.

Il contro-argomento per la tesi appena esposta è quello secondo cui la teoria propone un sistema — un sistema del tipo *GPS* — che incarna «una comunanza tra solutori di problemi» nella sua struttura di base. È quella struttura di base, quell'incarnazione della comunanza tra solutori di problemi, che rende relativamente facile scrivere programmi per risolvere problemi in una varietà di settori piuttosto disparati, per esempio la logica e la criptoaritmetica, in un sistema del tipo *GPS*. Ma, come hanno detto gli stessi Newell e Simon, qualunque comunanza di quel tipo, se non è attribuibile direttamente alla struttura di base, deve essere rappresentata o nel programma scritto in formalismo *GPS*, oppure nelle strutture di memoria. In realtà, tutte le versioni attuali di sistemi del tipo *GPS* hanno una struttura così assolutamente minima che l'informazione che bisogna dare loro (sotto forma di programmi e dati) per qualsiasi particolare compito deve essere dettagliata e specifica, vale a dire deve

definire quali siano gli operatori importanti, a quali oggetti essi possano venire applicati, che «differenza» facciano quando applicati all'oggetto appropriato, e così via. Come dicono Newell e Simon:

> Bisogna tenere nel giusto conto le limitazioni dell'accesso del *GPS* al mondo esterno. La parte iniziale delle istruzioni esplicite al *GPS* è stata acquisita dagli esseri umani molto tempo fa, quando essi hanno costruito il loro vocabolario generale. Questa [informazione] deve essere data a chiare lettere al *GPS* [23].

È proprio qui che la domanda viene spontanea. Perché la vera domanda è, che cosa succede all'uomo nella sua interezza quando egli mette insieme il suo vocabolario generale? In che modo la sua percezione di che cosa sia un «problema» viene influenzata dalle esperienze che sono parte integrante della sua acquisizione di un vocabolario? In che modo quelle esperienze influenzano la sua percezione di quali «oggetti», «operatori», «differenze», «obiettivi» ecc. sono importanti per un qualunque problema che egli si trovi di fronte? E così via. Nessuna teoria che eluda queste domande potrà mai essere una teoria della soluzione dei problemi da parte dell'uomo.

Il sogno dell'intellighenzia artificiale — espressione felice che il mondo deve al dottor Louis Fein — è, naturalmente, quello di creare «macchine che pensano, imparano, creano» e la cui abilità nel fare queste cose aumenti fino a che «il campo di problemi che sono in grado di trattare abbia estensione paragonabile a quella del campo cui è stata applicata la mente umana», come i dottori Newell e Simon annunciavano già nel 1958 [24]. Il loro libro venne pubblicato quattordici anni dopo e, come avevano promesso, «la capacità (delle macchine) nel fare queste cose aumentò rapidamente» anche se il «futuro visibile» allora sembra non essere ancora arrivato. Ma la visione è ancora chiara abbastanza. Ora, anzi, essi ci hanno detto qual è il trucco. Il sistema proposto, se gli si dà un tempo sufficiente (ma entro il futuro visibile), diventerà comprensivo di tutti i dettagli di una biblioteca di enciclopedie. Per poter diventare così comprensivo, tuttavia, anch'esso dovrà acquisire un vocabolario generale paragonabile a quello dominato da un uomo adulto; dovrà padroneggiare il linguaggio naturale e interiorizzare una base di conoscenze di estensione pari a quella dominata dalla mente umana. Una grande frazione della comunità che lavora sull'intelligenza artificiale si sta, in realtà, concentrando sul problema della comprensione del linguaggio naturale da parte del computer. È questo il problema che intendo discutere nel prossimo capitolo.

Per il momento, però, resta da chiedersi quale immagine dell'uomo come solutore di problemi possa generare — non dirò «giustificare» — la visione sbalorditiva che abbiamo presentato. Per rispondere, dobbiamo per prima cosa guardare a ciò che Newell e Simon intendono per «problema».

[23] *Ivi*, p. 855.
[24] Simon-Newell, *op. cit.*, p. 8.

Newell e Simon scrivono:

> Se noi forniamo una rappresentazione [di qual è il risultato desiderato, in quali condizioni, per mezzo di quali strumenti e operazioni, a partire da quali informazioni iniziali, e con quale accesso a risorse], e supponiamo che l'interpretazione [delle strutture simboliche che rappresentano questa informazione] sia implicita nel programma del sistema di elaborazione di dati per risolvere problemi, ecco che abbiamo definito un problema [25].

E allora, naturalmente, poiché «tutti gli uomini sono sistemi di elaborazione dell'informazione», si possono applicare a loro e ai loro affari i «metodi principali del *GPS*», cioè «l'analisi euristica mezzi-fini», il controllo di oggetti per stabilire se siano «indesiderati» (e quindi ancora da trasformare, con operatori, negli «oggetti desiderati»), e così via.

Si può obiettare che una simile caratterizzazione degli obiettivi dell'intelligenza artificiale è un gioco di parole, che sopravvaluta ingiustamente gli effettivi e molto più modesti obiettivi dell'*IA*, e che Newell e Simon e la comunità che lavora sull'*IA* in generale parlano soltanto di una certa classe di problemi tecnici, cui la precedente definizione può essere applicata, e per cui metodi del tipo *GPS* sono sicuramente appropriati. Ma il punto è proprio che la pervasione — si potrebbe anche dire perversione — del pensiero quotidiano da parte della metafora del computer ha trasformato ogni problema in un problema tecnico, per cui si pensa che i metodi descritti qui siano appropriati. Più avanti avrò qualcos'altro da dire su questo tema.

Per ora, basti osservare che H.A. Simon aveva scritto già nel 1960:

> Supponiamo che un particolare sviluppo tecnologico permetta l'automazione della stessa psichiatria, così che uno psichiatra possa fare il lavoro per cui prima ce ne volevano dieci.[...]: Questo esempio sembrerà del tutto immaginario soltanto a persone ignare di certi aspetti delle ricerche ora in corso sulla possibile automazione dei processi psichiatrici [26].

La ricerca cui si riferiva era quella appena incominciata da Kenneth Mark Colby, uno psicoanalista che scriveva:

> Avendo condotto molti esperimenti di laboratorio sulla libera associazione e avendo alle spalle anni di esperienza clinica con processi nevrotici, la mia speranza iniziale era di simulare insieme [!] il pensiero a libera associazione caratteristico di un processo nevrotico e i suoi cambiamenti sotto l'influenza degli interventi di uno psicoterapeuta [27].

[25] Newell-Simon, *Human Problem Solving*, p. 73.

[26] H.A. Simon, *The Shape of Automation* («La forma dell'automazione»), 1960, ristampato in Z.W. Pylyshyn (a cura di), *Perspectives on the Computer Revolution* («Opinioni sulla rivoluzione del computer»), Prentice-Hall, Englewood Cliffs, N.J. 1970, p. 413.

[27] K.M. Colby, *Simulation of Belief Systems*, («Simulazioni di sistemi di credenze»), cap. 6 in R.C. Schank-K.M. Colby (a cura di), *Computer Models of Thought and Language* («Modelli computerizzati di pensiero e linguaggio»), W.H. Freeman and Co., San Francisco, 1973, p. 257.

Il progetto — fortunatamente — fallì. Ma le parole di Simon dovevano echeggiare nelle orecchie del dottor Colby per altri sei anni, prima di emergere di nuovo dalla sua penna. Come già osservato, proprio il mio lavoro sul sistema *Eliza* riaccese il suo entusiasmo e lo spinse a scrivere i passaggi che ho citato in precedenza, ma che vale la pena di ripetere qui:

> Se il metodo (*Eliza*) si dimostra benefico, fornirà uno strumento terapeutico che potrà essere messo estesamente a disposizione di ospedali e centri psichiatrici che soffrono di carenza di terapeuti. |...| Con un sistema di computer si potrebbero trattare centinaia di pazienti all'ora [28].

Proprio come predetto da Simon, e oltre! Ovviamente, questa promessa euforica deriva proprio da una visione dell'uomo come sistema del tipo *GPS*. Nelle parole del dottor Colby,

> Un terapeuta umano può essere visto come un elaboratore di informazioni e un *decision maker* con un insieme di regole di decisione strettamente collegate a obiettivi di breve e di lungo termine. |...| In queste decisioni egli è guidato da rozze regole empiriche che gli dicono cosa è opportuno dire e non dire in certi contesti [29].

Il paziente, in altre parole, è un oggetto diverso da quello desiderato. Il compito del terapeuta è riconoscere la differenza, usando operatori scopri-differenze, e poi ridurla, usando operatori riduci-differenza, e così via. Questo è il suo «problema»! E questo è il punto a cui la metafora del computer ha portato qualcuno di noi.

[28] K.M. Colby-J.B. Watt-J.P. Gilbert, *A Computer Method of Psychotherapy: Preliminary Communication* («Un metodo computerizzato per la psicoterapia: comunicazione preliminare»), «The Journal of Nervous and Mental Disease«, vol. 142, 1966, nr 2, pp. 148-152.

[29] *Ivi*, p. 15.

7
IL COMPUTER E IL LINGUAGGIO NATURALE

Abbiamo distinto tre regimi nella ricerca sull'intelligenza artificiale: i cosiddetti regimi di prestazione, di simulazione e di teoria. Abbiamo notato, però, che la distinzione fra loro non è del tutto netta. Inoltre, abbiamo concluso che la parola «teoria», quando viene usata nell'espressione «regime di teoria», va presa in senso non proprio letterale. L'uso di idee ricavate dai computer e dal calcolo nei tentativi di capire la mente umana è ben più metaforico, mettiamo, dell'uso di idee meccanicistiche nella comprensione dell'universo fisico. Ma se lasciamo da parte la gran massa di lavoro sulla moderna scienza dei computer che tratta di argomenti teorici riguardanti il calcolo stesso (per esempio, la teoria degli automi finiti o la teoria della struttura dei linguaggi di programmazione) oppure dell'applicazione diretta dei computer a compiti precisi, indipendentemente dal fatto che l'esecuzione di tali compiti possa considerarsi un comportamento intelligente quando è svolta dagli esseri umani (per esempio, la soluzione di sistemi di equazioni differenziali, o il controllo computerizzato di certi processi chimici complessi), restiamo con un sottoinsieme della scienza dei computer in cui almeno uno degli obiettivi principali è l'imitazione dell'uomo da parte della macchina. Per gli scopi di questo capitolo, non risulterebbe utile sottolineare i diversi modi in cui il lavoro in questo campo può venire assegnato in modo predominante alla psicologia, alla linguistica o a chissà quale altra disciplina istituzionale. In quanto segue non insisterò perciò su queste distinzioni.

Due cose sono chiare: se vogliamo che una macchina faccia qualcosa, dobbiamo dirle di farlo, e la macchina deve poter capire ciò che le diciamo. Il modo più consueto per dire a un computer.cosa fare, almeno oggi, consiste nel dargli un particolare programma per il compito che abbiamo in mente e, naturalmente, i dati cui quel programma va applicato. Possiamo, per esempio, dargli un programma per calcolare radici quadrate e il numero 25, e aspettarci che ci restituisca il numero 5. Il computer «capisce» il programma per le radici quadrate, nel senso che sa interpretarlo esattamente nel modo che avevamo in mente quando l'abbiamo composto. Ma allora, un programma simile trasforma un computer in una macchina molto specializzata, una macchina per calcolare radici quadrate, e niente di più. Gli esseri umani, ammesso che siano macchine, sono macchine molto generali e, cosa più importante, capiscono comunicazio-

ni inviate in linguaggio naturale (per esempio, in italiano) che sono di gran lunga meno precise e più ambigue rispetto ai normali linguaggi di programmazione. Poiché l'obiettivo complessivo dell'*IA* è costruire macchine che «comprendano tutti i dettagli di una biblioteca di enciclopedie», bisogna naturalmente fare in modo che possano capire il linguaggio naturale. Ma anche lasciando da parte questi sogni, ci sono ragioni sia pratiche che scientifiche per lavorare sul problema del linguaggio naturale. Se si vuole che persone esterne al campo dei computer possano interagire con il computer in modo significativo, allora è necessario che esse imparino i linguaggi del computer, oppure che il computer impari il loro. Già oggi è più semplice insegnare al computer il linguaggio gergale di alcuni specialisti — per esempio certi medici, oppure ricercatori che lavorano sulle rocce lunari — piuttosto che insegnare a quegli specialisti il normale linguaggio dei computer. Alcuni scienziati del computer ritengono che le loro teorie sul linguaggio siano in un certo senso non del tutto legittime fintantoché rimangono quelle che la pubblica opinione spregevolmente chiama «pure teorie», cioè finché non si dimostri che esse possono essere convertite in modelli sotto forma di programmi per computer. D'altra parte molti linguisti, per esempio Noam Chomsky, credono che sia ancora necessario riflettere sul linguaggio quanto basta per tenerli occupati per un bel po', e che ogni sforzo per trasformare le loro teorie attuali in modelli per il computer sarebbe, se attuato dalle persone più qualificate, una deviazione dal compito principale. E giustamente essi non vedono alcuna ragione per spendere le loro energie nello studio del lavoro degli hacker.

Ma per il vero iniziato dell'intellighenzia artificiale, non è necessario affermare esplicitamente alcun motivo per lavorare sul problema della comprensione del linguaggio naturale da parte delle macchine. La capacità umana di manipolare simboli, la sua stessa capacità di pensare, sono inestricabilmente legate alle sue capacità linguistiche. Qualsiasi ri-creazione dell'uomo sotto forma di macchina deve perciò catturare questa, che è la più essenziale delle sue caratteristiche peculiari.

Non c'è, ovviamente, alcun *singolo* problema che possa venire ragionevolmente chiamato *il* problema del linguaggio naturale per i computer, così come non c'è un tale problema singolo per l'uomo. Ci sono invece molti problemi, che tutti hanno a che fare con il modo in cui un computer può essere messo in grado di capire messaggi che gli arrivano dal mondo esterno. Il problema della visione computerizzata, per esempio, è da molti punti di vista fondamentalmente analogo a quello della comprensione computerizzata del linguaggio naturale. Qualunque sia il modo in cui si fanno ricavare alla macchina informazioni dal suo ambiente, essa deve in qualche senso «capirle»; vale a dire, il computer deve in qualche modo saper estrarre il contenuto semantico dai messaggi che gli arrivano, in parte utilizzando la loro struttura puramente sintattica. Può sembrare strano, a prima vista, parlare della struttura sintattica di una scena visuale, e collegare il processo della sua comprensione a quello della comprensione di un testo in linguaggio naturale. Ma considerate l'immagine di un adulto e di un bambino su un'altalena. Noi comprendiamo certi aspetti di quella scena dalla sua forma, sebbene anche quella comprensione dipenda

dal fatto che abbiamo in precedenza adottato un certo schema concettuale, un insieme di convenzioni. Queste convenzioni sono sintattiche, nel senso che fungono da criteri che ci permettono di distinguere le immagini legittimamente ammissibili, per così dire, da quelle assurde. Le convenzioni figurative normalmente accettate, per esempio, farebbero rifiutare come non grammaticali quasi tutti i disegni di Escher. Tuttavia, noi comprendiamo l'immagine dell'altalena anche sulla base di spunti semantici. Sappiamo, per esempio, che la figura adulta, essendo in basso, è più pesante del bambino che sta seduto all'altro estremo della sbarra. E questa nozione ci deriva da qualcos'altro che non è la forma dell'immagine, in quanto coinvolge la nostra conoscenza privata di certi aspetti del mondo reale.

La comprensione del linguaggio — tanto da parte dell'uomo quanto della macchina — funziona anch'essa così. Tutti abbiamo certi criteri, una grammatica interiorizzata della lingua italiana, che ci permette di dire che la sequenza di parole «Il banco non l'ascia scampo» è «non grammaticale». Questo è un giudizio puramente sintattico. D'altra parte, ci rendiamo conto che la frase «Il banco non lascia scampo» è grammaticale, anche se possiamo avere qualche difficoltà per decidere cosa significhi, vale a dire come capirla. Diciamo che la capiamo soltanto quando abbiamo potuto costruire una storia in cui essa abbia senso, cioè quando possiamo indicare un qualche schema concettuale al cui interno la frase abbia un significato, magari anche «ovvio». Per esempio, in una storia in cui un giocatore d'azzardo ha puntato tutto su un suo sistema, e ha perso, la frase «Il banco non lascia scampo» ha un significato perfettamente ovvio, per lo meno per chi frequenta il casinò. Ancora una volta, è necessario focalizzare delle nozioni del mondo esterno, non soltanto per rendere la frase meno ambigua, ma proprio per poterle assegnare un qualunque significato.

Naturalmente, è molto più facile riuscire a controllare il problema della comprensione da parte della macchina del linguaggio naturale, piuttosto che il problema corrispondente per la visione, innanzitutto perché il linguaggio può essere rappresentato in forma scritta, cioè come sequenze di simboli appartenenti a un alfabeto molto ristretto. Inoltre, queste espressioni possono venir presentate al computer in serie sequenziale, proprio come vengono presentate al lettore-uomo. Possono anche venir memorizzate con fedeltà assoluta. Al contrario, il problema di che cosa costituisca un simbolo visivo, per quanto primitivo, già causa enormi difficoltà sia sintattiche che semantiche. Chi lavori sulla comprensione da parte della macchina di testi in italiano non si impegna a seguire alcuna particolare ipotesi, o strategia di ricerca, quando adotta certi simboli come primitivi, cioè come non analizzabili di per sé. Ma chi lavora sui problemi della visione avrà in pratica determinato delle componenti cruciali della sua strategia di ricerca nel momento in cui decide di adottare, mettiamo, spigoli e angoli come elementi del suo vocabolario primitivo. Inoltre, egli ha di fronte un problema formidabile proprio nel decidere quand'è che la sua macchina «vede» uno spigolo. Per questa ragione, e anche perché solo di recente si sono accoppiate telecamere ai computer, il lavoro sulla comprensione del linguaggio naturale da parte dei computer ha una storia molto più lunga, nella

ricerca sull'intelligenza artificiale, rispetto al lavoro sul problema della visione da parte del computer.

È successo molte volte nella storia del calcolo moderno che un progresso tecnologico nel campo dell'*hardware* o nella programmazione (*software*) abbia innescato una mania praticamente euforica. Quando quelli che all'epoca erano considerati computer di grandi dimensioni incominciarono a funzionare in modo più o meno affidabile, alcune persone altrimenti ragionevoli caddero vittime di quella che io chiamo sindrome dell'«Ora che (finalmente) abbiamo *X*, possiamo fare *Y*». Nella situazione cui mi riferisco, *X* erano quelli che all'epoca venivano considerati serbatoi di informazioni (memorie) molto grandi, mentre *Y* era la traduzione del linguaggio da parte della macchina. (Qui non farò dei nomi, dal momento che le persone contagiate allora sono ormai rinsavite da tempo).

La visione iniziale era, come avrebbe detto in seguito Robert K. Lindsay, «che si sarebbero potute produrre traduzioni di alta qualità con macchine dotate di regole sintattiche sufficientemente dettagliate, di un vasto dizionario e di velocità sufficiente per esaminare in ogni direzione il contesto di poche parole ambigue» [1]. I computer ancora oggi non producono «traduzioni di alta qualità». Tuttavia, un intransigente dogmatico della vecchia scuola — se ancora ve ne sono — potrebbe sostenere che non abbiamo ancora «regole sintattiche sufficientemente dettagliate» o «velocità sufficiente» per raggiungere il fine desiderato. Ma la vera domanda è se questa sufficienza sia possibile. Se si avesse un insieme qualsivoglia di regole sintattiche, dettagliate a piacere, e si potesse raggiungere qualsiasi velocità di calcolo, e un dizionario di qualsiasi dimensione, ciò basterebbe per ottenere traduzioni di alta qualità? Qualunque persona seria che lavori in questo campo ammette oggi che la risposta a questa domanda è semplicemente «no».

Una traduzione va vista come un processo che implica due componenti distinte, ma non del tutto separabili: bisogna capire il testo da tradurre e bisogna produrre il testo nel linguaggio-bersaglio, in cui si vuole tradurre. Per il nostro scopo qui, possiamo ignorare la seconda di queste componenti. Il problema si mostra quasi in tutta la sua complessità se consideriamo il caso in cui il linguaggio-bersaglio è lo stesso linguaggio-fonte, e trasformiamo così il problema della traduzione «semplicemente» in quello della parafrasi. Abbiamo visto che la comprensione anche di una sola frase può implicare tanto uno schema concettuale elaborato — per esempio, uno scenario che ha a che fare con casinò e giochi d'azzardo — e nozioni del mondo reale — per esempio, che cosa significa avere il banco, e così via. Ritorniamo a quella analogia. Supponiamo che la frase che abbiamo citato, «Il banco non lascia scampo», si fosse presentata nel primo capitolo di un romanzo giallo. La soluzione del delitto da parte dell'investigatore potrebbe dipendere dal fatto che egli arrivi a capire

[1] R.K. Lindsay, *Inferential Memory as the Basis of Machines which Understand Natural Language* («La memoria inferenziale come base di macchine in grado di comprendere il linguaggio naturale»), in E.A. Feigenbaum - J.Feldman (a cura di), *Computers and Thought* («Computer e pensiero»), McGraw-Hill, New York 1963, p. 218.

che questa frase si riferisce a una bisca. Ma gli spunti che portano a quell'interpretazione possono venire rivelati soltanto gradualmente, mettiamo uno per capitolo. Allora nessun uomo e nessun computer potrebbero capire, e quindi parafrasare, quella frase alla sua prima apparizione. Né sarebbe d'aiuto andare a leggere qualche parola che sta prima o dopo quella frase. Sia l'uomo che il computer dovrebbero leggere tutto il libro, tranne l'ultimo capitolo, per poter fare ciò che il detective fa alla fine, cioè capire la frase cruciale (supponiamo che l'ultimo capitolo del giallo serva soltanto a coloro che non hanno colto gli elementi che suggerivano come sciogliere il mistero, o comunque non sono riusciti a capire prima la soluzione). E anche allora, soltanto chi ha un'opportuna conoscenza del mondo potrebbe farlo.

Il riconoscimento del fatto che uno schema concettuale è essenziale per comprendere testi naturali venne sfruttato per la prima volta dai cosidetti «sistemi a domanda-risposta». B.F. Green e altri hanno scritto nel 1961 un sistema capace di capire, e rispondere, a domande sul baseball, per esempio [2]. Poteva capire la domanda «Dove ha giocato la tale squadra in luglio?» senza difficoltà perché, nel suo universo di discorso, parole potenzialmente problematiche come «squadra» e «giocare» potevano avere soltanto un unico significato. Sapeva rispondere, perché ciascuna domanda compresa in modo non ambiguo poteva facilmente essere trasformata in un piccolo programma per cercare informazioni pertinenti nella banca-dati del sistema. Il programma *Student* di Bobrow, benché molto più ambizioso, sfruttava lo stesso principio [3]. Era capace di risolvere problemi cosiddetti di algebra verbale come «Gianni ha un numero di pesci che è due volte il numero di criceti di Lisa. Se Lisa ha tre criceti, quanti pesci ha Gianni?». Anche qui, l'universo di discorso al cui interno il programma era destinato a operare determinava il modo in cui parole e frasi andavano capite, ricostruite (come formule algebriche) o altrimenti manipolate. Notate per esempio che, per «capire» il problema citato, non è affatto necessario «capire» parole come «pesci» e «criceti»; avrebbero potuto altrettanto bene essere «*X*» e «*Y*», rispettivamente. E la parola «ha» possiede una connotazione ben diversa da quella che avrebbe nella frase «Gianni ha il raffreddore». La specificazione di un universo di discorso molto vincolato semplifica enormemente il compito della comprensione; questo, naturalmente, vale anche per la comunicazione umana.

Ovviamente, in quasi tutte le situazioni realistiche la comprensione deve essere reciproca. Nel contesto della comunicazione uomo-macchina, vogliamo che la macchina ci capisca perché possa fare qualcosa per noi, ad esempio rispondere a una domanda, risolvere un problema matematico, o guidare un veicolo; azioni che noi, a nostra volta, speriamo di capire. Gli esempi citati non

[2] B.F. Green et al. *Baseball, an Automatic Question Answering System* («Baseball, un sistema automatico per rispondere alle domande»), in Feigenbaum -Feldman, *op. cit.*, pp. 207-216.

[3] D.G. Bobrow, *Natural Language Input for a Computer Problem-Solving System* («Dati di ingresso in linguaggio naturale per un sistema computerizzato per la risoluzione dei problemi»), in M. Minsky (a cura di), *Semantic Information Processing* («Elaborazione semantica dell'informazione»), The MIT Press, Cambridge, Mass. 1968, pp. 135-215.

gettano alcuna luce su questo aspetto della comunicazione uomo-macchina. Le risposte fornite dal programma *Baseball*, o da *Student*, semplicemente non hanno una portata interpretativa tale da essere problematici. Non si può immaginare di avere con loro una conversazione interessante. Tra le altre cose, fatto assai significativo, essi per parte loro non fanno domande.

Il primo programma che abbia illuminato quest'altra faccia del problema della comunicazione uomo-macchina è stato il mio *Eliza*[4]. Questo era un programma consistente per lo più in metodi generali per analizzare frasi e frammenti di frasi, identificando le cosiddette parole-chiave in un testo, riunendo frammenti in frasi, e così via. Esso non aveva, in altre parole, alcuno schema concettuale o universo di discorso incorporato. Quest'ultimo gli veniva fornito da un «copione». In un certo senso, *Eliza* era come un'attrice che conosca un insieme di tecniche di recitazione, ma che non abbia niente di suo da dire. Il copione, a sua volta, era un insieme di regole che consentivano all'attrice di improvvisare su un qualsiasi argomento.

Il primo copione approfondito che preparai per *Eliza* permetteva al programma di parodiare il comportamento di uno psicoterapeuta non direttivo in un colloquio psichiatrico preliminare. Scelsi questo copione perché mi permetteva di evitare temporaneamente il problema di dare al programma una banca-dati di nozioni del mondo reale. Dopo tutto, pensavo, uno psichiatra può ribattere all'affermazione del paziente «Mammà mi ha portato via l'orsacchiotto» dicendo «Mi parli ancora dei suoi genitori» senza dovere necessariamente saper nulla di orsacchiotti, per esempio. Per generare quella risposta, il programma doveva sapere che «Mammà» significa «madre» e che il «paziente» stava dicendo qualcosa su uno dei suoi genitori. Per la verità, il programma raggranellava anche altre informazioni dai messaggi del paziente, alcune delle quali potevano per esempio essere usate in risposte successive. Eppure, si poteva dire che «capisse» soltanto nell'accezione più debole possibile[5].

Ciònonostante, *Eliza* creò la straordinaria illusione di aver capito, nella mente di molte persone con cui conversò. Persone che sapevano perfettamente di stare conversando con una macchina dimenticarono ben presto questo fatto, proprio come degli spettatori a teatro, presi da un attacco di credulità, presto dimenticano che l'azione cui stanno assistendo non è «reale». A questa illusione si aggrappavano in modo particolarmente forte e tenacissimo quelle persone che sapevano poco o niente dei computer. Costoro spesso chiedevano di poter conversare con il sistema in privato, e dopo averlo fatto insistevano, nonostante le mie spiegazioni, nella convinzione che la macchina davvero li capisse.

[4] Weizenbaum, *Contextual Understanding by Computers* («La comprensione contestuale da parte dei computer»), in «Communications of the ACM», vol. 10, 1967, nr. 8, pp. 474-480.

Scelsi il nome «Eliza» perché, come l'Eliza del *Pigmalione* di G.B. Shaw, il programma poteva imparare a «parlare» sempre meglio anche se, come con Miss Doolittle, non era mai chiaro se diventasse o meno più intelligente.

[5] Cfr. tuttavia, il capitolo 3 a p. 81. Lí il programma *Eliza* ha un copione che gli permette di «capire» in un senso molto più forte.

Questo fenomeno è paragonabile alla convinzione — che molte persone hanno — che coloro che leggono la fortuna abbiano davvero delle intuizioni profonde, «conoscano il destino» e così via. Questa credenza non è una conclusione raggiunta dopo aver attentamente pesato prove favorevoli e contrarie. È piuttosto un'ipotesi che, nella mente di quanti ci credono, viene confermata dalle affermazioni dell'indovino. Come tale, svolge il ruolo del lampione dell'ubriaco di cui abbiamo parlato prima: non si permette alla luce di illuminare alcuna possibile prova contraria; e, anzi, tutto ciò che potrebbe esser considerato prova contraria da un osservatore disinteressato viene interpretato in un modo che elabora e rende più forte l'ipotesi.

Entro certi limiti, questo è un processo del tutto normale, e anche necessario. Nessuna «informazione» costituisce un dato, se non alla luce di una qualche ipotesi. Perciò, anche in una normale conversazione tra due persone, ciascun partecipante concentra parte di sé sul processo di comprensione dell'altro. Ciascuno dei due, in altre parole, ha un'ipotesi di lavoro, che di nuovo è uno schema concettuale, su che cosa l'altro è e sull'argomento della conversazione. Quest'ipotesi serve a predire cosa l'altro stia per dire e, cosa più importante, cosa intenda dire con ciò che sta per dire. La previsione funziona simultaneamente su parecchi livelli distinti. Al livello più basso, un ascoltatore si fa un'idea di quali saranno le prossime parole di colui che sta parlando, e completa per lui frasi ancora incompiute. Se, per esempio, su un ascensore si sente la voce all'altoparlante dire «Questo ascensore non ferma....», ci si aspetta che completi la frase con «al piano tale», ma non con «il secondo giovedì di ogni mese». Talvolta chi ascolta predice in modo errato, e pone riparo al danno che ne risulta soltanto molto più avanti nella conversazione, cioè quando ha di fronte l'evidenza schiacciante che deve aver «sentito male». Spesso, tuttavia, la predizione sbagliata viene falsificata prima che la frase in questione venga completata da chi parla. Chi ascolta allora si corregge al volo, e in modo praticamente inconsapevole.

A un livello molto più alto, ciascun partecipante porta nella conversazione un'immagine di chi sia l'altro per lui. Poiché è impossibile per chiunque conoscere completamente un'altra persona, quell'immagine consiste in parte di attribuzioni che si fanno all'identità dell'altro, attribuzioni che devono necessariamente essere basate su indizi ricavati da esperienze di vita indipendenti dei partecipanti. Il nostro riconoscimento di un'altra persona è così un atto di induzione su indizi che ci vengono forniti, in parte da lei e in parte dalla nostra ricostruzione del resto del mondo; è una sorta di generalizzazione. Siamo, in altri termini, tutti pieni di pregiudizi — nel senso di «giudicare prima» — sugli altri. E, come abbiamo notato, troviamo tutti difficile, o addirittura quasi impossibile, *accorgerci* di indizi che tendono a smentire i nostri giudizi (per non parlare dell'accettarli, e del permettere loro di diventare operativi).

È quindi facile capire perché certe persone che conversano con *Eliza* credano di essere capite, e si aggrappino a questa convinzione. Il «senso» e la continuità percepiti dalla persona che conversa con *Eliza* sono in gran parte forniti dalla persona stessa. Essa assegna a ciò che *Eliza* «dice» significati e inter-

pretazioni che confermano la sua ipotesi iniziale che il sistema capisca, proprio come potrebbe fare con ciò che gli dice un indovino. Tutto quello che *Eliza*, o l'indovino, devono fare è dare risposte che siano sufficientemente plausibili, e lascino abbastanza campo all'interpretazione, da rendere possibili quelle elaborazioni. E, siccome il soggetto non può stabilire i veri limiti delle capacità di *Eliza* (egli ha, dopotutto, solo un tempo limitato per giocarci, e il programma estrae da lui continuamente nuovo materiale), egli non può fare a meno di attribuirgli più potere di quanto ne abbia in realtà. Inoltre, egli sa che *Eliza* è stato costruito da un professore in un'università. Il programma è perciò avvolto nel mantello magico della scienza, e gli si possono attribuire tutti i ben noti poteri della scienza.

Il programma *Eliza*, in effetti, generava reazioni plausibili a quanto gli veniva detto. Per poterlo fare, doveva essere dotato anch'esso di un insieme di «attese». Queste erano codificate in tutti i copioni che gli venivano dati. A una persona che giocasse con *Eliza* in versione-psichiatra veniva chiesto di fare affermazioni del tipo di quelle che si fanno con uno psichiatra in una delle prime sedute. Gli veniva detto, in altre parole, quali erano le attese di *Eliza*. A un livello inferiore, il copione psichiatrico per *Eliza* era costituito in un modo tale da permettergli di fare predizioni locali su frasi e frammenti di testo, cioè di applicar loro ipotesi che esami successivi potessero confermare o smentire. Per esempio, il copione psichiatrico comprendeva l'ipotesi iniziale che un frammento della forma generale «tutti mi...», sebbene rappresentasse apertamente un messaggio sul rapporto tra il soggetto e «tutti», per esempio «tutti mi odiano», oppure su ciò che tutti fanno al soggetto, per esempio «tutti mi ridono dietro», si riferisse in modo celato e più importante a un recente incidente che avesse coinvolto il soggetto e una, o al massimo poche altre persone. La reazione di *Eliza* poteva perciò essere «Mi dica, chi Le ha detto di odiarLa, negli ultimi giorni?» oppure «Chi ha riso dietro le Sue spalle ultimamente?»

Ciò che distingue nettamente il lavoro attuale sulla comprensione del linguaggio naturale da parte della macchina da quello dei primi anni Sessanta è proprio il forte uso che attualmente si fa della predizione, sia al livello sintattico locale che, cosa più importante, al più ampio livello concettuale. Roger C. Schank, un giovane rappresentante di straordinaria bravura della scuola moderna, basa la sua teoria sull'idea cruciale che ogni espressione in linguaggio naturale sia una manifestazione, una codificazione, di una struttura concettuale sottostante. Capire un'espressione significa codificarla (Schank usa il termine tecnico *mapping*) nella propria struttura concettuale interiore.

> Qualsiasi coppia di espressioni di cui si possa dire che significano la stessa cosa, che siano nello stesso linguaggio o in linguaggi diversi, dovrebbero essere caratterizzate in un modo unico dalle strutture concettuali [...] La rappresentazione di questo contenuto concettuale, allora, deve essere in termini il più possibile interlinguali e neutri. [...] Ci [preoccuperemo di] trovare, ogni volta che viene detto qualcosa, una rappresentazione che spieghi il significato di quell'espressione in modo non ambiguo, e che possa essere trasformata a ritroso in quell'espressione, o in qualsiasi altra espressione che abbia lo stesso significato.

> Il punto importante è che, dietro ogni frase in un linguaggio, esiste almeno una concettualizzazione [6].

Quello che voglio sottolineare qui è che la teoria di Schank propone una struttura formale per la base concettuale che sta dietro alle espressioni linguistiche, che essa propone meccanismi specifici (algoritmi) per basare le predizioni su quelle strutture concettuali, e che propone delle regole formali per analizzare le espressioni in linguaggio naturale e per convertirle nelle basi concettuali. Tuttavia, Schank non crede che l'intera base di concezioni di un individuo possa essere estratta esplicitamente da lui. Crede soltanto che una simile struttura concettuale esista dentro ciascuno di noi e che, se potesse essere esplicitata, potrebbe in linea di principio venir rappresentata dal formalismo che egli propone. Una difficoltà, che Schank naturalmente riconosce, è che la struttura concettuale di ciascun individuo cambia in continuazione.

Nel discutere il ruolo che la struttura concettuale di una persona svolge nel modo in cui essa partecipa alle conversazioni, scrivevo nel mio lavoro del 1967 intitolato *Comprensione contestuale*:

> In alcuni settori della vita intellettuale dell'individuo, questa struttura può essere organizzata in modo altamente logico, almeno fino a un certo punto; per esempio, nel campo della sua professione. In settori più influenzati dalle emozioni, la struttura può venire organizzata in modo molto più impreciso, e può anche contenere molte contraddizioni. Quando una persona incomincia una conversazione, porta con sé la sua struttura concettuale come una specie di agenda.
> La struttura concettuale di una persona è il risultato della sua intera esperienza di vita. Tutte le persone hanno alcune esperienze formative in comune, per esempio sono tutte nate da una madre. Di conseguenza, qualche base per la comprensione tra due esseri umani esiste, già solo per il fatto di essere tali. Ma anche degli uomini che vivono nella stessa cultura troveranno difficoltà a capirsi, se la loro vita differisce in modo radicale. Poiché, in ultima analisi, ciascuna vita è unica, c'è un limite a quello che possiamo far capire a un'altra persona. Ciascuno di noi ha una *privacy* che preclude assolutamente la piena comunicazione delle nostre idee all'universo esterno a noi, e che quindi isola ciascuno di noi da ogni altro oggetto noetico nel mondo.
> Non ci può essere comprensione totale, né prova assolutamente affidabile della comprensione.
> Sapere con certezza che una persona ha capito ciò che le si è detto significa afferrare la sua intera struttura concettuale, e *questo* è equivalente a condividere la sua intera esperienza di vita. È proprio contro barriere di questo tipo che lottano gli artisti, e specialmente i poeti.
> Questo è un argomento da affrontare, se si vuole arrivare a un accordo su cosa significhi «comprensione» da parte della macchina. Ciò che il ragionamento precedente vuole mettere in chiaro è che si va troppo in là se si sostiene che una

[6] R. C. Schank, *Identifications of Conceptualizations Underlying Natural Language* («Identificazione di alcune concettualizzazioni che stanno dietro al linguaggio naturale») in R. Schank-K.M. Colby (a cura di), *Computer Models of Thought and Language* («Modelli computerizzati di pensiero e linguaggio»), W.H. Freeman and Co., San Francisco 1973, p. 191.

> macchina capirà una frase (o una sinfonia, o una poesia) soltanto se quella frase evocherà nella macchina le stesse immagini che erano presenti in chi parlava quando l'ha pronunciata. Perché, secondo questo criterio, nessun essere umano capisce alcun altro essere umano. Eppure, siamo d'accordo che gli uomini si capiscono tra loro *entro un margine di tolleranza accettabile*. La parola operativa qui è «accettabile», poiché implica uno *scopo*. Quando, perciò, parliamo della comprensione da parte di una macchina, parliamo di una comprensione limitata da qualche obiettivo. Chi afferma che ci sono idee che nessuna macchina riuscirà mai a capire, può al massimo voler dire che la macchina non capirà queste idee accettabilmente bene, poiché esse si riferiscono a obiettivi che, a suo giudizio, sono inappropriati rispetto alle macchine. Naturalmente, la macchina può ancora trattare quelle idee in modo simbolico, cioè in modi che riflettano — per quanto pallidamente — i modi in cui quegli obiettivi sono affrontati dagli organismi per cui sono appropriati [7].

Mi aspetterei che Schank, e con lui la maggior parte di coloro che oggi lavorano in questo campo, trovasse queste idee coerenti con le loro. Tuttavia, quando ho usato il termine «immagini» («quella frase evocherà nella macchina le stesse immagini»), termine che ora vedo come grosso modo corrispondente alle «strutture concettuali» di Schank, non avevo alcuna idea di come quelle immagini potessero essere rappresentate in un sistema formale.

Oggi esistono parecchi sistemi per la comprensione del linguaggio da parte del computer, che si basano su idee assai simili a quelle di Schank, benché vi si sia arrivati in modo indipendente, e vi siano delle differenze importanti. Uno dei migliori tra questi, e anche uno dei meglio conosciuti, è quello di Terry Winograd. Winograd, all'epoca studente di dottorato al Laboratorio di intelligenza artificiale del MIT, lavorava in stretta collaborazione con un gruppo che stava costruendo una macchina computerizzata «mano-occhio», cioè un sistema di computer che potesse «vedere» il suo ambiente per mezzo di una telecamera ad esso collegata, e che potesse manipolare gli oggetti del suo ambiente per mezzo di braccia e mani meccaniche controllate dal computer. Un problema che questo sistema risolse — per dare un esempio delle sue capacità generali — fu quello di copiare una struttura fatta di blocchi da costruzione per bambini, per esempio consistente di archi, ponti, e così via, sulla base di un modello della struttura desiderata che le era stato mostrato. Winograd si dedicò al compito di sviluppare un sistema per permettere agli uomini di dare istruzioni a questo robot, in linguaggio naturale, riferendosi strettamente al mondo-giocattolo dei blocchi da costruzione in cui esso doveva operare.

> All'interno di questo mini- mondo, possiamo dare al computer una conoscenza di tipo profondo. [...] Il robot sa manipolare blocchi-giocattolo su una tavola che contiene oggetti semplici come una scatola. Nel corso di un dialogo, gli si può chiedere di manipolare gli oggetti, facendo operazioni come costruire una colonna o mettere oggetti nella scatola. Gli si può chiedere quale sia la configu-

[7] J. Weinzenbaum, *op. cit.*, p. 476.

razione attuale dei blocchi sulla tavola, quali eventi siano accaduti durante la conversazione e, in modo limitato, quale sia il suo modo di ragionare. Gli si possono raccontare piccoli fatti, che vengono aggiunti alla sua memoria di conoscenze per essere usati nei ragionamenti successivi. La conversazione procede all'interno di uno schema dinamico, in cui il computer partecipa attivamente, facendo azioni per cambiare il suo mondo-giocattolo, e discutendone [8].

Quello che segue è un esempio di conversazione con il robot:

Uomo: trova un blocco più alto di quello che hai in mano e mettilo nella scatola.
Robot: suppongo che «mettilo» si riferisca al blocco più alto di quello che ho in mano.
Uomo: OK.
Robot: (la mano posa il blocco che teneva e ne prende uno più alto).
Uomo: che cosa contiene la scatola?
Robot: la piramide blu e la scatola blu.
Uomo: da che cosa è sostenuta la piramide?
Robot: dalla scatola.
|...|
Uomo: perché hai tolto il cubo (rosso)?
Robot: per metterlo su un grande cubo verde.
Uomo: e perché?
Robot: per incolonnare il cubo rosso, un grande blocco rosso e il grande cubo verde.
Uomo: e perché hai fatto tutto questo?
Robot: perché mi hai chiesto di farlo [9].

Il programma di Winograd è un risultato importante. Esso mostra come sia davvero possibile riempire una particolare visione di certi aspetti del linguaggio con dettagli sufficienti a fornire un modello di lavoro. Esso consiste fondamentalmente di descrizioni procedurali, cioè programmi, dei processi cognitivi, ed è così un'incarnazione di un particolare approccio psicologico alla comprensione del linguaggio come processo. D'altra parte, il programma ha dei seri difetti, alcuni dei quali sono notati da Winograd stesso:

(Non tratta) tutte le implicazioni del considerare il linguaggio come processo di comunicazione tra due persone intelligenti. Un uomo che usi il linguaggio è sempre impegnato in un processo di ricerca della comprensione del mondo circostante, compresa la persona con cui sta parlando. Costruisce attivamente modelli e ipotesi, e ne fa uso nel processo di comprensione del linguaggio.

[8] T. Winograd, *A Procedural Model of Language Understanding* («Un modello procedurale della comprensione del linguaggio»), in Schank-Colby, *op. cit.*, p. 154.
[9] *Ivi*, pp. 155-163.

(Poiché il sistema) conserva memoria di quando le cose sono state citate, può controllare una possibile interpretazione di una domanda per vedere se chi l'ha fatta possa rispondere da sé in base alle sue frasi precedenti. Se è così, suppone che probabilmente egli intenda qualcos'altro. Potremmo dire che ciò contenga due tipi di conoscenze: in primo luogo, suppone che una persona faccia domande per avere informazioni che non ha già e, in secondo luogo, ha un modello molto primitivo di quali informazioni egli abbia, sulla base di ciò che ha detto. Una visione realistica del linguaggio deve avere un modello complesso di questo tipo, e l'euristica del nostro sistema tocca appena un pezzetto minuscolo delle conoscenze rilevanti [10].

Bisogna davvero dire che questa dimostrazione di umiltà è molto consolante, specialmente perché viene dall'interno del clero dell'intellighenzia artificiale, ed è quindi praticamente unica. Purtroppo, non va abbastanza a fondo. Perché quello che Winograd ha fatto — anzi, quello che tutto il lavoro sull'*IA* ha fatto finora — è costruire una macchina che svolge certi compiti specifici, proprio come, mettiamo, gli artigiani del XVII secolo costruivano macchine per tenere il conto del tempo, per lanciare palle di ferro a distanze notevoli, e così via. Questi artigiani si sarebbero dolorosamente ingannati se si fossero lasciati portare dai loro successi alla conclusione di essere vicini a una comprensione teorica generale dell'universo, o anche alla conclusione di aver dimostrato, siccome le loro macchine funzionavano, la validità dell'idea che le leggi dell'universo siano esprimibili in termini matematici. La *hubris* dell'intellighenzia artificiale si mostra proprio nella sua ricorrente adesione a queste idee sbagliate sulle macchine che è riuscita a costruire. Né l'umiltà di Winograd, né quella di alcun altro ricercatore di *IA* arrivano fino ad ammetterlo.

Newell, Simon, Schank e Winograd confondono semplicemente la natura dei problemi che credono di stare «risolvendo». Come se fossero artigiani arretrati del XVII secolo, essi presentano «teorie generali » che sono in realtà soltanto slogan euristici praticamente vuoti, e poi sostengono di aver verificato queste «teorie» costruendo modelli che in effetti eseguono qualche compito, ma in un modo che non fornisce alcuna intuizione di principi generali. Questo fallimento è intrinseco, poiché essi hanno mancato di riconoscere che, per poter fare quello che sostengono di fare, essi devono scoprire e formulare principi generali di potenza maggiore di quella intrinseca nell'osservazione — o anche nella dimostrazione — che sia possibile formulare leggi sotto forma di programmi per computer. L'effetto più importante, e di più vasta portata, di questo insuccesso è che i ricercatori nel campo dell'*IA* si illudono continuamente col credere che la ragione per cui qualsiasi particolare sistema non si è avvicinato a realizzare la grandiosa visione dell'*IA* vada sempre cercata nei limiti del programma di quel particolare sistema. Così, per esempio, Winograd riconosce che il suo sistema si avvale soltanto di una minuscola parte delle conoscenze necessarie. La conoscenza di cui egli parla è la conoscenza di «fatti» che è disponibile agli esseri umani. Ma il guaio del suo approccio è che la sua

[10] *Ivi*, p. 185.

euristica non esprime alcun principio generale interessante. Inoltre, simili principî non si possono scoprire semplicemente estendendo il dominio di un sistema in modo tale da permettergli di acquisire maggiori conoscenze sul mondo. Neppure il più intelligente costruttore di orologi del XVII secolo avrebbe mai scoperto le leggi di Newton, semplicemente costruendo orologi sempre più elaborati e intricati!

L'intelligenza artificiale, come abbiamo documentato, si è posta come obiettivo la costruzione di macchine il cui dominio di pensiero abbia estensione paragonabile a quello dell'umanità stessa (non preoccupiamoci, per ora, del fatto che a questo si debba arrivare nel «futuro visibile» o meno). E le teorie che devono puntellare questo trionfo dell'*IA* devono valere anche per l'uomo nella sua interezza. Chiaramente, allora, le limitazioni cui sono soggetti i sistemi oggi esistenti — alcune delle quali sono confessate da Winograd con sincera umiltà — sono viste dalla comunità dell'*IA* nel suo complesso come difficoltà solo temporanee, che possono essere superate (nel futuro visibile, secondo Newell e Simon). Ci sono allora due domande che vanno infine affrontate: primo, le basi concettuali che stanno dietro alla comprensione linguistica sono completamente formalizzabili, se pure in linea di principio, come Schank suggerisce e come credono i più tra quanti lavorano nell'*IA*? Secondo, esistono idee che, come ho suggerito, «nessuna macchina potrà mai capire, perché si riferiscono a obiettivi non appropriati per le macchine»?.

Queste due domande sono di enorme importanza. Vanno al cuore del problema dell'esistenza di differenze essenziali tra uomo e macchina. Ed è opportuno farle nel contesto di una discussione del problema della comprensione del linguaggio naturale da parte delle macchine, perché è nel suo linguaggio — in fin dei conti — che l'uomo manifesta la sua intelligenza e, come credono alcuni, la sua identità unica di uomo. Le due domande devono anche essere fatte insieme. Esse sono inestricabilmente legate l'una all'altra. Perché se il complesso di una esperienza umana, e la struttura concettuale cui questa dà origine, non possono essere formalizzati, allora ci sono davvero obiettivi propriamente umani, che non sono appropriati per le macchine. E se dovessimo concludere (come intendo fare) che obiettivi simili ci sono davvero, allora potremmo anche dire qualcosa su ciò che si dovrebbe e non si dovrebbe far fare alle macchine.

Il fatto stesso che queste domande siano diventate importanti è indicativo della profondità cui la metafora dell'elaborazione di informazioni ha penetrato la mentalità accademica e quella popolare. Perché, quando facciamo l'inventario, scopriamo rapidamente quanto poco si sia di fatto realizzato fino a ora. Il libro di Newell e Simon, *Human Problem Solving*, parla in dettaglio soltanto di tre problemi: la criptoaritmetica, la risoluzione di problemi nel calcolo logico più elementare, e gli scacchi. I risultati ottenuti dalle macchine computerizzate «mano-occhio» al MIT e alla Stanford University (per esempio, la costruzione di strutture di blocchi a partire da modelli mostrati alla macchina, l'avvitare dadi su bulloni) sono giustamente salutati come trionfi da quanti comprendono l'incredibile complessità dei problemi che hanno dovuto essere risolti per arrivarci. E ci sono stati altri trionfi di grandezza analoga. Ma

il fatto stesso che risultati simili meritino l'applauso testimonia di quanto profondamente primitiva sia la nostra attuale conoscenza della mente umana. George A. Miller, come i moderni linguisti del computer, parla anch'egli delle strutture concettuali che stanno dietro al pensiero e al linguaggio umano, ma egli dice: «Fare finta che sappiamo come impartire queste strutture concettuali complesse a una qualunque macchina, al giorno d'oggi, è semplicemente assurdo» [11]. Winograd parla davvero a nome dell'intero settore dell'intelligenza artificiale, che i suoi colleghi lo riconoscano o meno, quando dice che i nostri sistemi hanno sfiorato soltanto una minuscola parte delle conoscenze necessarie.

Ma proprio come sarebbe stato scorretto, nel XVII secolo, sostenere l'impossibilità di mettere un oggetto fatto dall'uomo in orbita intorno alla Terra, solo per il fatto che nessuno a quell'epoca aveva la più lontana idea di come fare, così sarebbe sbagliato oggi fare affermazioni di impossibilità sui computer soltanto in base alla nostra attuale ignoranza. È pertinente, tuttavia, specialmente per fare il punto sulla nostra situazione attuale, esaminare la potenza delle teorie che abbiamo discusso. Sono, per esempio, newtoniane nella vastità della loro portata inferenziale?

Che contributo si dà quando si afferma che «esiste una base concettuale che è interlinguale, su cui le strutture linguistiche di un dato linguaggio si proiettano durante il processo della comprensione e da cui tali strutture vengono create durante la generazione (di espressioni linguistiche)»? [12]. Assolutamente nulla. Perché il termine «base concettuale» potrebbe benissimo venir sostituito dalla parola «qualcosa». E chi potrebbe mettere in discussione l'affermazione così trasformata? Il contributo di Schank, come quelli di altri che ora coltivano lo stesso fertile terreno, sta nel tentativo di fornire una rappresentazione formale della base concettuale. Egli intende dirci in tutti i dettagli che cosa sia quel qualcosa «che sta dietro a tutti i linguaggi naturali», e come funzioni, sia nella generazione che nella comprensione delle espressioni linguistiche. Eppure, Schank non dimostra che il suo schema sia qualcosa di più di una collezione di regole euristiche che casualmente funzionano su particolari classi di esempi. Il problema scientifico cruciale sarebbe quello di costruire un programma finito, che assegni strutture concettuali appropriate all'infinita varietà di frasi che possono presentarsi nel linguaggio naturale. Quel problema rimane intatto come sempre. Immaginate una macchina per fare somme che addizioni correttamente certi numeri ma non altri, e di cui non sappiamo nemmeno caratterizzare la classe dei numeri che sa sommare. Questa non la chiameremmo certo «meccanizzazione dell'aritmetica». E in che forma ci verranno mostrate le strutture concettuali che Schank ipotizza, le operazioni su di loro, i collegamenti tra loro, e così via? Ma sotto forma di programmi per computer, naturalmente.

[11] G.A. Miller, *Language, Learning, and Models of the Mind* («Linguaggio, apprendimento e modelli della mente»), manoscritto non pubblicato, giugno 1972.

[12] R.C. Schank, *Conceptual Dependency: A Theory of Natural Language Understanding* («La dipendenza concettuale: una teoria della comprensione del linguaggio naturale»), in «Cognitive Psychology», 1972, nr. 3, pp. 553-554.

> Ciò che è necessario, ed è mancato finora, è una teoria coerente di come gli uomini capiscono il linguaggio naturale, relativa non a sottoparti particolari di quel problema, ma al problema nel suo complesso. La teoria [...] è anche progettata per essere una base per programmi per computer che capiscano il linguaggio naturale. [...] Ciò che verrà discusso è la teoria di un programma di questo tipo. [...] Speriamo di riuscire a costruire un programma che possa imparare, come fa un bambino, a fare ciò che abbiamo descritto in questo lavoro, invece di venire imboccato con l'enorme quantità di informazioni necessarie. Per poter fare questo, potrebbe essere necessario attendere di avere un sistema mano-occhio automatico efficace, e un elaboratore di immagini [13].

Qui iniziamo a vedere la confluenza del lavoro sulla risoluzione dei problemi, che abbiamo discusso in precedenza, con il lavoro sulla comprensione del linguaggio naturale da parte delle macchine. Una frase è un «oggetto dato», la struttura concettuale che ne è il significato è l'«oggetto desiderato», l'obiettivo è trasformare il primo nel secondo, e i mezzi sono quelli forniti dal programma per la comprensione. Ma la cosa più importante nei due casi è che le teorie siano convertibili in programmi per il computer.

Può essere possibile, seguendo le procedure di Schank, costruire una struttura concettuale che corrisponda al significato della frase «Verrai a cena con me stasera?». Ma è difficile vedere — e so che questo non significa che sia impossibile — come degli schemi come quelli di Schank potrebbero mai capire che quella frase rappresenta il disperato desiderio d'amore di un giovane timido. Anche se un computer potesse simulare sensazioni di disperazione e di amore, è possibile che un computer *sia* disperato e *provi* amore? Può allora il computer capire disperazione e amore? Nei limiti in cui queste domande sono legittime — e sono limiti molto angusti — la risposta è «no». E se questa è la risposta, allora il senso in cui anche il più potente sistema del tipo di Schank «capisce» è più o meno debole come il senso in cui *Eliza* «capiva».

Nella migliore delle ipotesi, quello che vediamo qui è un altro esempio di ricerca delle chiavi da parte dell'ubriacone. Una teoria pretende di descrivere le strutture concettuali che stanno dietro a tutta la comprensione del linguaggio da parte dell'uomo. Ma le uniche strutture concettuali che ammette come legittime sono quelle che possono essere rappresentate sotto forma di strutture di dati manipolabili dal computer. Poi, si annuncia semplicemente che queste ultime costituiscono tutte le strutture concettuali che stanno alla base di tutto il pensiero umano. Dato un simile programma, cioè una tale restrizione del significato della parola «tutto», dovrebbe davvero essere possibile dimostrare che la teoria spieghi «tutto» il comportamento linguistico dell'uomo!

Una teoria, naturalmente, è essa stessa uno schema concettuale. E così, determina che cosa deve e cosa non deve contare come fatto. Le teorie — sarebbe forse meglio dire le metafore originarie — che hanno ipnotizzato l'intellighenzia artificiale, e anche vasti settori dell'opinione pubblica, hanno stabili-

[13] *Ivi*, pp. 553-629.

to molto tempo fa che la vita è ciò che si può calcolare, e nient'altro. Come ha detto il professor John McCarthy, capo del Laboratorio di intelligenza artificiale alla Stanford University, «L'unica ragione per cui non siamo ancora riusciti a formalizzare tutti gli aspetti del mondo reale è che ci è mancato un sistema di calcolo logico sufficientemente potente. Attualmente sto lavorando su questo problema».

A volte, quando i miei figli erano ancora piccoli, mia moglie e io stavamo a guardarli mentre dormivano. Ci parlavamo in silenzio, recitando una scena antica come l'umanità. Come Ionesco disse al suo diario: «Non *tutto* è indicibile a parole; soltanto la verità vivente».

8
L'INTELLIGENZA ARTIFICIALE

Quando Roger Schank espresse la speranza di poter costruire un programma capace di imparare come un bambino, egli echeggiava parole pronunciate da H.A. Simon più di dieci anni prima:

> Se il *GPS* è una teoria di come una macchina si può sollevare a un livello di intelligenza superiore, o di come gli uomini imparano il linguaggio, allora facciamo che si sollevi, e facciamole imparare il linguaggio. È del tutto opportuno imporre quest'obbligo. [...] Non solo a nome di me stesso, ma dell'intero gruppo di persone che lavorano in questo campo, accetto quest'obbligo e spero che uno di noi tra non molto produrrà il programma richiesto [1].

Sia Simon che Schank hanno così dato espressione alla visione più profonda e grandiosa che motiva il lavoro sull'intelligenza artificiale, cioè niente di meno che la costruzione di una macchina sul modello dell'uomo, di un robot che abbia un'infanzia, impari il linguaggio come fa un bambino, ricavi la sua conoscenza del mondo «sentendolo» attraverso organi suoi propri, e contempli infine l'intero campo del pensiero umano. (Vale la pena di osservare, anche se solo per inciso per il momento, che se questo sogno si dovesse realizzare avremmo una macchina che comprende il linguaggio, ma non ancora una teoria della comprensione del linguaggio in quanto tale, perché osservare una macchina «che impara come un bambino» non costituisce di per sé una comprensione del processo di acquisizione del linguaggio).

Il fatto che questo programma possa o meno essere realizzato dipende dal fatto che l'uomo sia soltanto una specie del genere «sistemi di elaborazione dell'informazione», oppure qualcosa di più. Io sosterrò che il pensiero popolare e quello scientifico sono stati dominati da una nozione di intelligenza del tutto semplicistica, e che questa nozione è, in parte, responsabile di aver permesso alle grandiose e perverse visioni dell'intelligenza artificiale di crescere. Sosterrò che un organismo è definito, in larga misura, dai problemi che af-

[1] In M. Greenberger (a cura di), *Management and the Computer of the Future* («La gestione aziendale e il computer del futuro»), The MIT Press, Cambridge, Mass. 1962, p. 123.

fronta. L'uomo affronta problemi che non si potranno mai far affrontare a una macchina. L'uomo non è una macchina. Sosterrò che, benché sia assolutamente certo che l'uomo elabora informazioni, non necessariamente lo fa nel modo in cui lo fanno i computer. Uomo e computer non sono specie diverse dello stesso genere.

Pochi concetti «scientifici» hanno generato confusione, nel modo di pensare tanto degli scienziati quanto dell'opinione pubblica come quello di «quoziente d'intelligenza» o «QI». L'idea che l'intelligenza possa essere misurata quantitativamente lungo una semplice scala lineare ha causato un danno indicibile alla nostra società in generale, e all'istruzione in particolare. Ha alimentato, per esempio, l'enorme movimento per i test attitudinali negli Stati Uniti, che influenza fortemente il corso degli studi di milioni di giovani, e così il livello dei diplomi che essi riescono ad ottenere. In pratica, esso determina il «successo» che potranno avere nella vita perché, almeno negli Stati Uniti, le possibilità di «avere successo» sono aperte per lo più soltanto a coloro che hanno le credenziali adeguate, vale a dire lauree universitarie, diplomi professionali ecc.

Quando gli educatori moderni sostengono che i test d'intelligenza misurano la capacità di un soggetto di far bene a scuola, essenzialmente intendono dire che questi test «predicono» la capacità di un soggetto di superare esami di tipo accademico. Questa capacità, naturalmente, porta al diploma e quindi al «successo». In conseguenza, qualsiasi correlazione tra il risultato di quei test e il «successo» delle persone, nel senso in cui questo termine è inteso nella società in generale, deve essere necessariamente un artefatto della procedura stessa dei test: il test è diventato un criterio per ciò con cui deve essere correlato! «Gli psicologi dovrebbero vergognarsi di promuovere una visione dell'intelligenza generale che ha creato un simile programma di test» [2].

Ciò che mi interessa qui è il fatto che la mitologia che circonda i test QI ha portato a una convinzione molto diffusa, e profondamente fuorviante, secondo cui l'intelligenza è una sorta di attributo permanente, inalterabile, indipendente dalla cultura degli individui (un po' come, diciamo, il colore degli occhi), e che per di più potrebbe anche venire trasmessa geneticamente di generazione in generazione.

Il guaio del test QI non è che sia del tutto falso, quanto che è incompleto. Misura certe capacità intellettuali che larghi settori, politicamente dominanti, delle società occidentali europee hanno elevato a vera e propria essenza del valore umano, e quindi a *conditio sine qua non* del successo. È incompleto in due sensi: primo, perché non tiene conto del fatto che la creatività umana non dipende solo dall'intelletto, ma anche e in modo cruciale da un'interazione tra intelletto e altre modalità di pensiero, come intuito e saggezza; secondo, perché caratterizza l'intelligenza come fenomeno misurabile linearmente, che esiste indipendentemente da alcun sistema di riferimento.

Einstein ci ha insegnato che l'idea di moto è priva di significato di per sé, che possiamo parlare sensatamente soltanto del moto relativo di un ogget-

[2] D.C. McClelland, *Testing for Competence rather than for «Intelligence»* («Fare test di competenza piuttosto che di "intelligenza" »), *American Psychologist*, vol. 28, 1973, nr. 1, pp. 1-14.

to rispetto a un qualche sistema di riferimento, non di moto *assoluto* di un oggetto. Quando, nel linguaggio quotidiano, diciamo che un treno si muove, intendiamo dire che si muove rispetto a un certo punto fisso sulla terra. Nei discorsi normali non è necessario sottolinearlo, perché la terra (o il nostro corpo) è per noi una sorta di sistema di riferimento convenzionale, che è supposto e compreso come tale implicitamente nella maggior parte delle conversazioni informali. Ma un fisico che parli da scienziato non può essere così impreciso. Le sue equazioni del moto devono contenere dei termini che specificano il sistema di coordinate rispetto a cui il moto descritto ha luogo.

Lo stesso vale per l'intelligenza. Quello di intelligenza è un concetto privo di significato, di per sé. Richiede un sistema di riferimento, la precisazione di un dominio di pensiero e azione, per poter diventare significativo. La ragione per cui questa necessità non ci si manifesta quando parliamo di intelligenza nei discorsi quotidiani, è che il sistema di riferimento in questione — vale a dire, la nostra particolare struttura sociale e culturale, con i suoi peculiari domini di pensiero e azione — è così connaturato in noi che implicitamente lo diamo per scontato. Ma la nostra cultura e il nostro ambiente sociale non sono in realtà né universali né assoluti. È perciò necessario, ogni qualvolta usiamo il termine «intelligenza» parlando da scienziati o da educatori, che esplicitiamo il dominio di pensiero e azione che rende comprensibile quel termine.

La nostra vita quotidiana dimostra più che a sufficienza che l'intelligenza si manifesta soltanto relativamente a precisi contesti sociali e culturali. La madre più ignorante dal punto di vista scolastico, che non sa scrivere una sola frase grammaticalmente corretta nella sua madrelingua — come in realtà è il caso anche per molti accademici —, ragiona continuamente in modo molto raffinato e intelligente sui problemi della sua famiglia. Studiosi eminenti confessano di non avere il tipo di intelligenza necessario per capire l'algebra del liceo. Il genio riconosciuto talvolta è stupido nell'organizzare la sua vita privata. I computer sono capaci di prodigiose *performance* intellettuali, come battere campioni di scacchi al loro gioco e risolvere enormi sistemi di equazioni, ma non sanno cambiare il pannolino a un bambino. Come si fa a paragonare tra loro queste intelligenze? Non si possono paragonare.

Eppure, diverse forme dell'idea che l'intelligenza sia misurabile lungo una scala assoluta, e quindi che le intelligenze siano paragonabili, hanno penetrato in profondità il pensiero odierno. Quest'idea è responsabile, almeno in parte, di molti sterili dibattiti per stabilire se sia possibile, in linea di principio, costruire computer più intelligenti dell'uomo. Perfino uno psicologo moderato e ragionevole come George A. Miller scivola di tanto in tanto su questa buccia di banana, come quando dice: «Sono molto ottimista su quello che sarà il risultato del lavoro sulla soluzione di problemi intellettuali da parte delle macchine. Entro la fine della nostra generazione, può darsi che le macchine ci superino in intelligenza generale»[3].

L'identificazione dell'intelligenza con il QI ha gravemente distorto la questione, principalmente matematica, di cosa i computer possano e non possano

[3] Grenberger, *op. cit.*, p. 118.

fare, trasformandola nella questione priva di senso di «quanta» intelligenza si possa, di nuovo «in linea di principio», dare a un computer. E, naturalmente, la spericolata antropomorfizzazione del computer così comune oggi, specialmente tra l'intellighenzia artificiale, ben si accorda con queste visioni semplicistiche dell'intelligenza. Questa unione di una metafora illegittima con un'idea mal formulata dà origine — e sembra dare legittimità, in apparenza — a proposizioni perverse come, ad esempio, quella per cui un computer può essere programmato per diventare uno psicoterapeuta efficace.

Un tempo ho sperato che sarebbe stato possibile dimostrare che c'è un limite superiore all'intelligenza che le macchine possono raggiungere, proprio come Claude Shannon, il fondatore della moderna teoria dell'informazione, aveva dimostrato che c'è un limite superiore alla quantità di informazione che un singolo canale di informazione può trasmettere. Shannon dimostrò che, per esempio, uno specifico cavo telefonico può trasportare al massimo un certo numero di conversazioni telefoniche per volta. Tuttavia, prima di poter anche solo formulare in modo sensato questo risultato per cui ora è giustamente famoso, egli dovette trovare un modo per quantificare l'informazione. Altrimenti, come avrebbe potuto parlare della capacità da parte di un canale di maneggiare «tutta questa» informazione ma non «di più»? Anzi, il suo stesso progetto di una misura dell'informazione costituisce un importante contributo alla scienza moderna (posto, naturalmente, che egli fondò anche una teoria rigorosa al cui interno la sua misura gioca una parte decisiva). Oggi mi è chiaro che, dal momento che possiamo parlare di intelligenza soltanto in particolari domini di pensiero e azione, e siccome questi domini stessi non sono misurabili, non possiamo avere alcuna misura «alla Shannon» dell'intelligenza, né quindi alcun teorema del tipo che speravo. In parole povere: possiamo esprimere il desiderio, e perfino l'opinione, che ci sia un limite all'intelligenza cui le macchine possono arrivare, ma non abbiamo modo di dargli un significato preciso, né certamente di dimostrarlo.

Forse che la nostra incapacità di calcolare un limite superiore all'intelligenza delle macchine fornisce la base per la conclusione «ottimistica» che «forse le macchine ci supereranno in intelligenza generale» (oppure per la stessa conclusione vista però come «pessimistica»)? [4]. No. Ne ricaviamo invece l'insegnamento che qualsiasi discussione che tenda a confermare, o a negare, quella conclusione è per sua natura mal posta e perciò sterile.

Queste considerazioni gettano ancora più luce su una questione cui ho alluso nel capitolo 7 (p. 171), dove parlavo di «obiettivi che sono inappropriati per le macchine». Molti obietterebbero che, tanto per cominciare, non è ragionevole parlare delle macchine come se avessero degli obiettivi. Ma questa sottigliezza retorica, se presa sul serio, non fa che rinvigorire la domanda, perché ignora il fatto che la gente, all'atto pratico, delega davvero responsabilità ai computer, e dà loro obiettivi e scopi.

[4] L'ottimista dice: «Questo è il migliore dei mondi possibili!». Il pessimista ribatte: «È vero».

La domanda che sto cercando di affrontare qui è: «Quali possono essere gli obiettivi e scopi umani che non è bene delegare ai computer?». Possiamo progettare un pilota automatico, e delegargli il compito di far volare un aeroplano lungo una rotta predeterminata. Questa sembra una cosa appropriata da far fare alle macchine. È anche tecnicamente possibile costruire un sistema di computer che intervisti i pazienti che chiedono aiuto a un consultorio psichiatrico, e ne produca profili psichiatrici completi di tabelle, grafici e commenti in linguaggio naturale. Il problema non è se una cosa simile *possa* essere realizzata, bensì se sia appropriato delegare a una macchina questa funzione finora svolta da esseri umani.

L'intellighenzia artificiale sostiene, come abbiamo visto, che non vi è dominio del pensiero umano su cui le macchine non possano spaziare. Essi danno per scontato che le macchine possano concepire pensieri come quelli che uno psichiatra ha in mente quando è impegnato col suo paziente. Essi sostengono che considerazioni di costo ed efficienza impongono di delegare alle macchine quel tipo di responsabilità. Come mi disse una volta il professor John McCarthy durante un dibattito, «Cosa sa un giudice che non si possa dire a un computer?». La sua risposta alla domanda — che è in realtà la nostra stessa domanda, soltanto in una forma diversa — è naturalmente «Nulla». Ed è, come egli sostenne più avanti, perfettamente appropriato che l'intelligenza artificiale cerchi di costruire macchine capaci di produrre decisioni giuridiche.

L'affermazione secondo cui giudici e psichiatri non sanno nulla che non si possa dire ai computer, segue dall'affermazione molto più generale fatta propria dall'intellighenzia artificiale, quella cioè per cui non vi è proprio nulla che l'uomo sappia che non possa, almeno in linea di principio, essere in qualche modo reso accessibile ai computer.

Non tutti gli scienziati del computer sono ancora così ingenui da credere — come si imputava loro una volta — che la conoscenza consista semplicemente in una qualche organizzazione dei «fatti». I vari programmi per la comprensione del linguaggio e per la visione, per esempio, memorizzano parte della loro conoscenza sotto forma di affermazioni, cioè assiomi e teoremi, e parte sotto forma di processi. Anzi, mentre progettano ed eseguono alcune delle loro procedure complesse, questi programmi compongono dei sottoprogrammi, cioè generano nuovi processi, che non erano stati introdotti esplicitamente dai programmatori umani. Alcuni sistemi computerizzati esistenti, in particolare le cosiddette «macchine mano-occhio», accumulano conoscenza «sentendo» direttamente il loro ambiente. Queste macchine quindi possono venire a conoscere, oltre alle cose che vengono loro dette esplicitamente, anche cose che scoprono interagendo con il mondo. Infine, è possibile insegnare ai computer certe abilità pratiche, per esempio come tenere in equilibrio un manico di scopa per una sua estremità, mostrando loro come si fa, anche se l'istruttore stesso non è capace di tradurrre in parole il modo per riuscirci. Il fatto dunque — ed *è* un fatto — che gli uomini sappiano cose che non possono comunicare sotto forma di linguaggio scritto o parlato, non è di per sé sufficiente a stabilire che esistono conoscenze che i computer non potranno mai acquisire.

Ma, poiché temo che la mia «ammissione» che i computer hanno il potere di acquisire conoscenze in molti modi diversi possa far fraintendere ciò che ho inteso dire, lasciatemi esprimere la mia posizione in modo molto chiaro.

Per prima cosa (la meno importante), la capacità anche del più avanzato tra i sistemi di computer oggi esistenti di acquisire informazioni, con mezzi che non siano quello che Schank chiama «essere imboccati», è ancora estremamente limitata. La potenza dei metodi euristici esistenti per estrarre conoscenze da testi in linguaggio naturale, anche direttamente «messi in bocca» al computer, si basa precariamente, per usare le parole di Winograd, su «una porzione minuscola delle conoscenze necessarie». È semplicemente assurdo credere che un qualsiasi sistema di computer oggi esistente possa arrivare a conoscere in un modo qualunque ciò che, mettiamo, un bimbo di due anni sa sui blocchi da costruzione per bambini.

In secondo luogo, non è affatto ovvio che tutte le conoscenze umane siano codificabili in «strutture informatiche», per quanto complesse. Un uomo può sapere, per esempio, che tipo di impatto emotivo avrà il toccare la mano di un'altra persona, sia sull'altro che su di sé. L'acquisizione di quella conoscenza non è certo una funzione del solo cervello; non può essere semplicemente un processo in cui una struttura d'informazione viene trasmessa da una qualche sorgente nel mondo a una qualche destinazione nel cervello. Le conoscenze coinvolte sono in parte cinestetiche; per acquisirle, è necessario avere una mano, per dire proprio il minimo. Ci sono, in altre parole, delle cose che gli uomini sanno in virtù del fatto di avere un corpo umano. Nessun organismo che non abbia un corpo umano può sapere queste cose nello stesso modo in cui le sa l'uomo. Qualsiasi loro rappresentazione simbolica deve tralasciare certe informazioni, che sono essenziali per alcuni scopi umani.

Terzo, e l'esempio del toccare la mano può tornare utile anche qui, ci sono cose che gli uomini vengono a sapere soltanto come conseguenza dell'essere stati trattati come esseri umani da altri esseri umani. Tra un attimo ritornerò su questo.

Quarto e ultimo punto, anche quel tipo di conoscenze che superficialmente sembra comunicabile da un essere umano a un altro usando il solo linguaggio, non è in realtà così completamente comunicabile. Claude Shannon mostrò che, anche nella teoria astratta dell'informazione, il «contenuto d'informazione» di un messaggio non è funzione del solo messaggio, bensì dipende in modo cruciale dallo stato delle conoscenze, dalle attese di chi riceve. Il messaggio «Arrivo coll'aereo delle 7, ciao amore, Filippo» ha un contenuto di informazione diverso per la moglie di Filippo, che sapeva che sarebbe arrivato ma non sapeva con quale aeroplano, da quello che avrebbe per una ragazza che Filippo non lo aspettava proprio, ed è sorpresa dalla sua dichiarazione d'amore.

Il linguaggio umano, nel suo uso pratico, è infinitamente più problematico di quei suoi aspetti che possono venir trattati dalla teoria dell'informazione, naturalmente. Ma perfino l'esempio che ho citato dimostra che il linguaggio coinvolge le storie di coloro che lo usano, quindi la storia della società, anzi di tutta l'umanità in generale. E il linguaggio nell'uso umano non è solamente funzionale nel modo in cui sono funzionali i linguaggi per il computer.

Non identifica cose e parole soltanto con obiettivi immediati da raggiungere, o con oggetti da trasformare. L'uso umano del linguaggio rivela la memoria umana. E questa è un oggetto ben diverso dal serbatoio di informazioni del computer, che è stato antropomorfizzato chiamandolo «memoria». La nostra memoria dà origine a speranze e paure, per esempio. È difficile capire cosa potrebbe significare dire che un computer «spera».

Queste considerazioni toccano non soltanto certe limitazioni tecniche dei computer, ma anche il problema centrale di cosa significhi essere un essere umano, e cosa essere un computer.

Io accetto l'idea che un sistema computerizzato moderno sia sufficientemente complesso e autonomo da permetterci di parlarne come di un organismo. Dato che esso può percepire e influenzare il suo ambiente, concederò perfino che possa — in senso estremamente limitato — essere «socializzato», vale a dire modificato dalle sue esperienze con il suo mondo. Concedo anche che un robot opportunamente costruito può essere portato a sviluppare un certo senso di se stesso, e che può, per esempio, imparare a distinguere le parti proprie dagli oggetti esterni, che gli si può insegnare a dare priorità alla protezione delle sue parti, di fronte a un possibile danno fisico, piuttosto che alla protezione di oggetti esterni, e che esso può formarsi un modello di se stesso che, in un certo senso, potrebbe essere considerato una sorta di autocoscienza. Quando dico, perciò, che sono disponibile a considerare un simile robot come un «organismo», dichiaro la mia disponibilità a considerarlo una sorta di animale. E ho già ammesso che non vedo il modo di porre un limite al grado di intelligenza che un simile organismo potrebbe raggiungere, almeno in linea di principio.

Scrivo queste disposizioni, come le chiamerebbe un avvocato, non perché io creda che quello che qualunque osservatore ragionevole chiamerebbe «robot socializzato» verrà costruito nel «futuro visibile» — infatti non ci credo —, bensì per evitare l'esercizio non necessario, interminabile e in fin dei conti sterile di compilare un catalogo di ciò che i computer saranno e non saranno in grado di fare, qui e ora come nel futuro. Quest'esercizio ci distoglierebbe dal problema principale, vale a dire, se vi siano obiettivi che non è appropriato assegnare alle macchine.

Se tanto l'uomo quanto le macchine sono adattabili all'ambiente, socializzabili, allora dobbiamo chiederci in che cosa la socializzazione dell'uomo debba essere necessariamente diversa da quella della macchina. La risposta, naturalmente, è così ovvia da far sembrare ridicolo — se non osceno — lo stesso fatto di porre la domanda. Il fatto che questo argomento debba essere affrontato è, già di per sé, un segno della follia dei nostri tempi.

Ogni organismo viene socializzato dal processo di affrontare i problemi che ha davanti. Proprio le caratteristiche biologiche che differenziano una specie da un'altra determinano anche il fatto che quella specie affronterà problemi diversi da quelli di un'altra. Ciascuna specie, anche solo per questa ragione, verrà socializzata in modo diverso. Il bambino, come hanno osservato in molti, nasce prematuramente, cioè trovandosi completamente indifeso. Eppure il bambino ha necessità biologiche che, se si vuole che sopravviva, devono essere

soddisfatte da altri. Anzi, molti studi sugli orfani hanno mostrato che è necessario soddisfare altri bisogni, non solo quelli fisici elementari; un neonato morirà se viene nutrito e pulito, ma non coccolato e accarezzato fin dall'inizio della sua vita: se, in altre parole, non viene trattato come essere umano da altri esseri umani [5].

Una catastrofe, per usare l'espressione di Erik Erikson, che ogni essere umano deve affrontare è la sua ricapitolazione personale della storia biblica del paradiso terrestre. Per qualche tempo, il bambino chiede e riceve gratificazione per ogni suo bisogno, senza che gli venga chiesto nulla in cambio. Poi, spesso dopo che il bambino ha sviluppato i primi denti e ha morso il seno che l'ha nutrito, l'unità tra lui e la madre si rompe. Erikson ritiene che questo dramma umano universale costituisca il contributo ontogenetico alla saga biblica del Giardino dell'Eden. Questo periodo è così importante nella vita del bambino che

> una improvvisa perdita di amore materno, senza adeguata sostituzione, in questo momento può portare [in altre condizioni aggravanti] a un'acuta depressione infantile, oppure a uno stato di afflizione moderata ma cronica, che può lasciare un tono depressivo in tutto il resto della vita. Ma anche nelle circostanze più favorevoli, questa fase lascia un residuo di senso del male e del destino avverso, e una nostalgia universale per un paradiso perduto.
> (Questi stadi iniziali) allora, formano nel bambino le sorgenti del senso fondamentale di fiducia e di sfiducia che rimangono per tutta la vita le sorgenti autogene tanto della speranza primitiva quanto del senso di malasorte [6].

Così comincia la ricostruzione immaginaria del mondo da parte dell'individuo. E questo mondo, come ho detto in precedenza, è il ricettacolo della sua soggettività, lo stimolatore della sua coscienza, e in ultima analisi il costruttore delle forze apparentemente esterne che dovrà affrontare per tutta la vita.

> Quando il raggio di consapevolezza, coordinazione e sensibilità del bambino si espande, egli incontra i modelli educativi della sua cultura, e impara così le modalità fondamentali dell'esistenza umana, ciascuna in modo significativo tanto dal punto di vista personale quanto culturale. [...] *Prendere* [...] significa ricevere e accettare ciò che viene dato. Questa è la prima modalità sociale che si impara nella vita; e suona più semplice di quello che è. Perché l'organismo neonato, che è instabile e si muove a tentoni, impara questa modalità soltanto quando impara a regolare i suoi sistemi organici in accordo con il modo in cui l'ambiente materno coordina i suoi metodi di cura del bambino [...].
> La situazione complessiva ottimale implicata dalla prontezza del bimbo nel prendere ciò che gli viene dato è la sua reciproca regolazione con una madre che gli consenta di sviluppare e coordinare i suoi modi di prendere, mentre essa sviluppa e coordina i suoi modi di dare. [...] La bocca e il capezzolo sembrano essere

[5] Vedere in particolare il lavoro di R.A. Spitz, *Hospitalism*, in *Psychoanalytic Study of the Child*, vol. 1, 1945.

[6] E. Erikson, *Infanzia e società*, Armando, Roma 1982, pp. 79-80.

soltanto i centri di una generale aura di calore e reciprocità, e generano piacere e rilassamento non solo in questi organi focali, ma in entrambi gli organismi nel loro complesso. La reciprocità del rilassamento così generato è di primaria importanza per la prima esperienza di «diversità amichevole». Si può dire [...] che nel *prendere* in questo modo *ciò che viene dato,* e nell'imparare a *far fare a qualcuno* ciò che vuole gli sia fatto, il bambino sviluppa anche il fondamento di egoismo necessario per *riuscire a diventare* uno che dà [7].

Ciò che queste parole di Erikson mettono in chiaro è che le fasi iniziali e cruciali della socializzazione umana implicano e avviluppano la totalità di due organismi, il bimbo e sua madre, in una reciprocità inseparabile di servizio ai loro più profondi bisogni biologici ed emotivi. E da questa problematica riunificazione di madre e bambino — problematica perché implica inevitabilmente il trauma della separazione — emergono le fondamenta della conoscenza umana di cosa significhi dare e ricevere, fidarsi e non fidarsi, essere un amico e avere un amico, avere un senso di speranza e un senso di disperazione.

Prima, parlando delle teorie (p. 129), ho detto che non è mai possibile capire in modo completo e conclusivo alcun termine di una teoria. Possiamo dire lo stesso delle parole in generale, specialmente di parole come fiducia, amicizia, speranza e derivate. Erikson ci insegna che queste parole derivano il loro significato da esperienze umane universali, primitive, e che ogni comprensione di quelle parole deve sempre essere fondamentalmente metaforica. Questa verità profonda ci informa anche del fatto che l'intera comprensione umana del mondo, essendo mediata dal suo linguaggio, deve essere limitata sempre e necessariamente da descrizioni metaforiche. E siccome il bambino «incontra i modelli educativi della sua cultura», come dice Erikson, «e impara così le modalità fondamentali dell'esistenza umana, ciascuna in modo significativo tanto dal punto di vista personale quanto culturale», ciascuna cultura, anzi ciascun individuo in una cultura, comprende parole e linguaggio, e quindi il mondo, in un mondo culturalmente e personalmente idiosincratico.

Potrei continuare a descrivere gli stadi successivi della socializzazione dell'individuo umano, gli effetti della scuola, del matrimonio, del carcere, della guerra, di odii e amori, i sensi di vergogna e colpa che differiscono così radicalmente tra le diverse culture umane, e così via. Ma ciò non servirebbe a convincere chi a questo punto non sia già convinto che qualsiasi «comprensione» di cui si possa parlare a proposito di un computer, e quindi qualsiasi «intelligenza» che gli possa venire attribuita, possono avere soltanto un legame molto vago con la comprensione e l'intelligenza umana. Noi, tuttavia, concludiamo che per quanta intelligenza i computer possano raggiungere, ora o in futuro, la loro dovrà sempre essere un'intelligenza *estranea* ai problemi e agli interessi genuinamente umani.

Eppure, l'ala estrema o intransigente dell'intellighenzia artificiale insisterà nel sostenere che l'uomo nella interezza, per usare ancora una volta l'espressione di Simon, è dopo tutto un elaboratore di informazioni, e che una teoria

[7] *Ivi,* pp. 75-76.

dell'uomo basata sull'elaborazione dell'informazione deve perciò essere adeguata per rendere conto del suo comportamento in tutti i suoi aspetti. Possiamo essere d'accordo con la premessa maggiore, senza necessariamente dedurne la conclusione indicata. Abbiamo già osservato che una parte dell'informazione che l'uomo «elabora» è cinestetica, che essa è «memorizzata» nei suoi muscoli e nelle sue articolazioni. Semplicemente, non è affatto chiaro che tale informazione, e l'elaborazione a essa associata, possa essere rappresentata sotto forma di programmi e strutture di dati per il computer.

Si può, naturalmente, ribattere che è possibile in linea di principio per un computer simulare l'intera rete di cellule che costituiscono il corpo umano. Ma questo introdurrebbe una teoria dell'elaborazione dei dati completamente diversa da quelle proposte finora. Inoltre, una simulazione di quel tipo produrrebbe come risultato un «comportamento» soltanto in un tempo così incredibilmente lungo, che nessun robot costruito su questi principi potrebbe mai interagire con gli esseri umani. Infine, non sembra esserci alcuna probabilità che l'umanità impari abbastanza sulla neurofisiologia, per parecchi secoli a venire, da acquisire le basi intellettuali per il progetto di una macchina simile. Possiamo perciò accantonare queste obiezioni.

C'è, però, ancora un'altra ipotesi che i modellisti dell'uomo come elaboratore di informazioni fanno e che può essere falsa, e la cui negazione indebolisce assai il loro programma: che esista una e una sola classe di processi informatici, e che ogni membro di quella classe sia riducibile al tipo di processo informatico esemplificato da sistemi come il *GPS* e i formalismi alla Schank per la comprensione del linguaggio. Eppure ogni essere umano ha l'impressione di pensare per intuito, sensazioni e altri mezzi non formali, almeno tanto quanto «in modo sistematico», cioè con mezzi come la logica. In giro si continuano a sentire domande come «Può un computer avere idee originali? Può comporre una metafora, o una sinfonia, o una poesia?». È come se la saggezza popolare conoscesse la distinzione tra il pensiero del computer e il tipo di pensiero cui si dedicano normalmente le persone. Gli adepti dell'intellighenzia artificiale, naturalmente, non credono ci debba essere distinzione. Sorridono e rispondono: «non dimostrato».

Durante lo scorso decennio, tuttavia, hanno incominciato ad accumularsi delle prove neurologiche che suggeriscono che ci possa essere una base scientifica per la saggezza popolare[8]. Da molto tempo si sa che il cervello umano è composto da due cosiddetti emisferi che, almeno superficialmente, sembrano identici. Queste due metà, che chiameremo ES (emisfero sinistro) ed ED (destro), hanno però funzioni ben distinte. Nelle persone che si servono per lo più della mano destra — e, per semplicità, possiamo restringere a queste la nostra discussione — si può dire che l' ES, grosso modo, controlla la metà destra del corpo, e l'ED la metà sinistra (in realtà, la connessione è un po' più

[8] Per un rapporto sui risultati ottenuti in questo settore, vedere R.E. Ornstein, *La psicologia della coscienza*, Angeli, Milano 1978. Il capitolo 3 è particolarmente pertinente per la presente discussione. È scritto in forma molto semplice. La bibliografia che cita apre una finestra sull'intero settore di ricerca.

complessa, in particolare quella tra le due metà del cervello e gli occhi, ma non scenderò qui in simili dettagli). Cosa importantissima, sembra che le due metà del cervello abbiano due modalità di pensiero ben distinte. L'ES pensa, per così dire, in modo ordinato, sequenziale, che potremmo chiamare logico. L'ED, d'altra parte, sembra pensare in termini di immagini olistiche. L'elaborazione del linguaggio sembra essere centrata quasi esclusivamente nell'ES, per esempio, mentre l'ED è profondamente coinvolto in compiti come l'orientamento nello spazio e la produzione e l'apprezzamento della musica.

Le diverse funzioni dei due emisferi del cervello furono mostrate per la prima volta, e in modo spettacolare, da pazienti affetti da forme estremamente gravi di epilessia, cui le due metà del cervello vennero separate chirurgicamente. Nelle persone normali, i due emisferi sono collegati da una parte del cervello chiamata *corpus callosum*. Quando questo viene tagliato, non è più possibile alcuna comunicazione tra le due metà. Si è trovato che, quando le mani di un paziente «a cervello diviso» venivano nascoste alla sua vista, e gli veniva data — mettiamo — una matita nella mano sinistra, egli non sapeva *dire* cosa gli era stato dato, ma sapeva *mostrare* che si trattava di una matita, facendone uno schizzo oppure scegliendo l'immagine di una matita tra una serie di immagini di oggetti diversi. Tuttavia, quando l'esperimento veniva ripetuto affidando la matita alla mano destra, allora egli sapeva dire che era una matita, ma non ne sapeva produrre o riconoscere l'immagine grafica. Nella prima situazione, l'ED riceveva l'«immagine» della matita ed era in grado di codificarla in rappresentazioni illustrate, ma non in strutture linguistiche. Nel secondo caso, era l'ES a ricevere l'«immagine» della matita, ed esso era in grado di codificarla in modo linguistico, ma non illustrativo.

Ci sono anche molte indicazioni, al cui proposito non scenderò qui nei dettagli, del fatto che l'ED sia sostanzialmente la sede dell'intuito, e che pensi in modo indipendente dall'ES. Un modo per descrivere il pensiero intuitivo è dire che, benché esso sia logico, usa per formulare giudizi degli standard di evidenza molto diversi da quelli che di solito associamo al pensiero logico. Nella conversazione quotidiana, per esempio, quando diciamo che due cose sono la stessa cosa intendiamo dire che esse sono identiche da quasi tutti i punti di vista; lo standard di evidenza che richiediamo per giustificare un giudizio simile è estremamente stringente. Ma quando formuliamo una metafora, per esempio dicendo che i cinesi all'estero sono gli ebrei d'Oriente, asseriamo che due cose sono la stessa cosa, ma in un senso molto diverso. Le metafore sono semplicemente non logiche; se le si prende alla lettera, sono chiaramente assurde. L'ED, in altre parole, ha criteri di assurdità ben diversi da quelli del logico ES.

La storia della creatività umana è piena di vicende di artisti e scienziati i quali, dopo aver lavorato sodo e a lungo su un problema difficile, decidono consapevolmente di «dimenticarlo»; in realtà, di trasferirlo al loro ED. Dopo qualche tempo, spesso in modo molto improvviso e del tutto inaspettato, la soluzione al problema si annuncia loro in forma quasi completa. L'ED sembra esser stato capace di superare i problemi logici e sistematici più spinosi, e io tenderei a supporre che questo sia avvenuto grazie a un allentamento degli standard di pensiero rigorosi dell'ES. In virtù degli standard più liberi che adot-

ta, l'ED è stato magari capace di progettare esperimenti concettuali che l'ES non avrebbe semplicemente potuto concepire, a causa del suo rigore. L'ED può così azzeccare delle soluzioni che possono poi, naturalmente, venire riformulate in termini strettamente logici dall'ES. Possiamo ipotizzare che nei bambini il canale di comunicazione tra le due metà del cervello sia molto ampio; vale a dire, che messaggi passino molto liberamente dall'una all'altra. Questo potrebbe spiegare perché i bambini siano così incredibilmente fantasiosi; per loro, per fare un esempio, una scatola di sigari può essere ora un'automobile, e tra un momento una casa. Negli adulti, il canale è stato gravemente ristretto; se la ragione stia nell'educazione, o in processi fisiologici di maturazione, o in entrambi, io non lo so. Ma quel canale è chiaramente più aperto quando si sogna. Posso anche ipotizzare che la psicoanalisi, a parte la sua funzione come psicoterapia, alleni la gente a usare il canale. In analisi si impara, per usare la felice espressione coniata da Theodor Reik, ad ascoltare con il terzo orecchio, a fare cioè attenzione a ciò che sta «dicendo» l'inconscio. Forse, le varie discipline meditative servono allo stesso scopo.

Queste sono chiaramente delle ipotesi, da cui non abbiamo il diritto di trarre alcuna conclusione sul modo in cui gli uomini, oppure i computer, elaborano le informazioni. Anche come una pura possibilità, tuttavia, esse sollevano un'obiezione seria rispetto all'universalità del modo di elaborare informazioni che di solito associamo al pensiero logico e ai programmi per il computer.

Che l'emisfero destro del cervello sia, a grandi linee, la «sede dell'intuito», è un'ipotesi che sembra essere sostenuta da un numero crescente di prove. Né i filosofi né gli psicologi sono stati ancora convinti, dalle testimonianze esistenti, quanto basta per incorporare fiduciosamente questa ipotesi nelle loro teorie della mente. Ma c'è qualcosa di sicuramente stabilito: i due emisferi del cervello umano pensano in modo indipendente uno dall'altro; essi pensano simultaneamente; e pensano in modi diversi uno dall'altro. Inoltre, siamo in grado di dire qualcosa a proposito di questi due modi diversi.

Il grande matematico Henri Poincaré scrisse nel suo famoso saggio *La creazione matematica*[9]:

> L'io conscio è strettamente limitato, e come per l'io subliminale noi non conosciamo le sue limitazioni [...]. I calcoli vanno fatti nel [...] periodo di lavoro conscio, quello che segue l'ispirazione, quello in cui si verificano i risultati dell'ispirazione e se ne deducono le conseguenze. Le regole di questi calcoli sono rigide e complicate. Essi richiedono disciplina, attenzione, volontà e perciò consapevolezza. Nell'io subliminale, al contrario, regna quella che potrei chiamare libertà, se questo nome si può dare alla semplice assenza di disciplina [...]. I fenomeni inconsci privilegiati, quelli che possono diventare consci, sono quelli che, direttamente o indirettamente, influenzano in modo più profondo la nostra sensibili-

[9] Ristampato in *The World od Mathematics* («Il mondo della matematica»), Simon and Schuster,New York 1956, vol. IV, pp. 2041-2050. Questo importante saggio merita un salto in biblioteca, così come l'insieme dei volumi in cui compare.

tà emotiva. |...|. Il ruolo di questo lavoro inconscio nell'invenzione matematica mi sembra incontestabile, e se ne troverebbero tracce in altri casi ove è meno evidente.

Naturalmente Poincaré, scrivendo all'inizio del ventesimo secolo, non sapeva nulla delle scoperte dei ricercatori sul cervello attivi oggi. E salteremmo a una conclusione se identificassimo quelli che egli chiama l'io conscio e l'io subliminale con gli emisferi sinistro e destro del cervello, rispettivamente. Ma la nostra affermazione, qui, è che ci sono due regimi diversi del pensiero umano, che operano in modo indipendente e simultaneo. E le parole di Poincaré confortano questa affermazione.

Uno scienziato molto rispettato che è al lavoro oggi, lo psicologo Jerome Bruner, scrive su questo stesso argomento da un punto di vista leggermente diverso (ricordate che la mano destra corrisponde all'emisfero sinistro, e la sinistra all'emisfero destro, o «intuitivo»):

Da psicologo che usa la mano destra, ho studiato diligentemente per quindici anni i processi cognitivi: come acquisiamo, tratteniamo e trasformiamo conoscenze del mondo in cui ciascuno di noi vive; un modo in parte «esterno», in parte «interno» a noi. Gli strumenti che ho usato sono stati quelli dello psicologo scientifico che studia la percezione, la memoria, l'apprendimento, il pensiero e (figlio del mio tempo) ho svolto la mia indagine sui topi di laboratorio oltre che sugli esseri umani. A volte, anzi, ho adottato il ruolo del clinico e ho svolto terapia con i bambini. |...| Ci sono state occasioni in cui, un po' scoraggiato dalle complessità della psicologia della conoscenza, ho cercato una via di scampo attraverso la neurofisiologia, per scoprire che il neurofisiologo può essere d'aiuto soltanto nella misura in cui gli sappiamo porre delle domande psicologiche intelligenti.

Una cosa è divenuta sempre più chiara durante la caccia alla natura della conoscenza. Ed è che l'apparato tradizionale dello psicologo — tanto i suoi strumenti d'indagine quanto gli attrezzi concettuali che usa nell'interpretazione dei suoi dati — lascia inesplorato un possibile approccio. È un approccio la cui moneta di scambio sembra essere la metafora fornita dalla mano sinistra. È un sistema che coltiva presentimenti felici e intuizioni «fortunate», che viene mescolato in pozione fino a diventare attività di collegamento dal poeta e dal negromante, che guardano a lato anziché direttamente all'obiettivo. I loro presentimenti e le loro intuizioni generano una grammatica loro propria: ricerca di connessioni, suggerimento di analogie, tessitura libera di idee in un tentativo di tela. |...|

Anche (lo psicologo) va estesamente e metaforicamente alla ricerca dei suoi presentimenti. Egli legge romanzi, osserva quadri, ne dipinge perfino, è colpito dalla potenza del mito, osserva gli altri uomini con intuito e meraviglia. Nel far questo, egli agisce solo a metà tempo da vero e proprio psicologo, catalogando casi su casi secondo i criteri derivati dalle sue ipotesi. Come gli altri uomini, egli osserva il palcoscenico umano con tutta la sensibilità che riesce a mettere insieme, nella speranza di approfondire la sua comprensione. Se è fortunato, o ha sottile intuito psicologico, di tanto in tanto gli verranno delle illuminazioni, prodotto combinatorio della sua attività metaforica. Se non avrà paura di questi prodotti della sua stessa soggettività, si spingerà fino ad addomesticare le metafore che hanno prodotto l'illuminazione, addomesticarle nel senso di farle passare dalla mano sinistra alla destra, convertendole in nozioni che possano essere messe alla

prova. La mia impressione, ricavata osservando me stesso e i miei colleghi, è che presagi metaforici vengano continuamente forgiati fino a diventare ipotesi verificabili [10].

Questa, naturalmente, è anche la mia impressione. Qui Bruner parla esplicitamente della mano sinistra — cioè dell'emisfero destro del cervello — come artistica, intuitiva, e così via, e della mano destra — l'emisfero sinistro del cervello — come «apparato tradizionale dello psicologo», e parla dell'inadeguatezza dei suoi «attrezzi concettuali».

Dalla testimonianza di centinaia di persone creative, oltre che dalla nostra personale introspezione, apprendiamo che l'atto creativo dell'uomo implica sempre l'interpretazione conscia di messaggi provenienti dall'inconscio, il passaggio delle idee dalla mano sinistra alla destra, per usare l'espressione di Bruner.

L'inconscio è, naturalmente, inconscio. È come un mare ribollente e tempestoso dentro di noi. Le sue onde bagnano le frontiere della nostra coscienza. E ciò che apprendiamo da esso o su di esso, lo costruiamo a partire da illazioni che facciamo sul significato delle onde lunghe e dei cavalloni, dei frangenti e delle increspature che slavano i margini della nostra coscienza. Di tanto in tanto ci addentriamo un po' di più nella risacca, come quando siamo nello stato di *trance* semi-ipnotica che separa il sonno dallo stato vigile. Ma in quei casi incontriamo soltanto il caos. Le nostre modalità di pensiero sono confuse al massimo. E se ci costringiamo a svegliarci non sappiamo dire, non sappiamo tradurre o trasformare in modalità linguistiche, quello che avevamo pensato.

L'indubbia realtà di questa confusione, quando sia aggiunta a tutti gli altri segnali disponibili cui ho accennato, non rafforza forse l'ipotesi — del tutto plausibile — che le forme di informazione manipolate nell'emisfero destro del cervello, insieme ai processi informatici corrispondenti, siano semplicemente diverse da quelle dell'emisfero sinistro? E non può darsi che noi possiamo, in linea di principio, arrivare a conoscere queste strane forme e questi strani processi informatici soltanto in termini che sono sostanzialmente irrilevanti per il genere di comprensione di cui andiamo in cerca? Quando, in un futuro lontano, arriveremo a conoscere in dettaglio il modo in cui il cervello funziona al livello neurofisiologico, saremo naturalmente in grado di dare un rendiconto, in ultima analisi riduzionista, del funzionamento dell'emisfero destro. Ma questa non sarebbe una comprensione nel senso che intendiamo qui, non più di quanto una conoscenza dettagliata del comportamento elettrico di un computer in funzione sia, o anche porti a, una comprensione del programma che il computer sta eseguendo. D'altro canto, un resoconto di livello superiore sul funzionamento dell'emisfero destro può sempre lasciarsi sfuggire le sue caratteristiche più essenziali, vale a dire quelle che lo differenziano dal funzionamento dell'emisfero sinistro. Perché siamo costretti, dalle nostre modalità di pensiero dell'ES, a interpretare sempre i messaggi che arrivano dall'ED in termini propri dell'ES.

[10] J. Bruner, *Il conoscere*, Armando, Roma 1982, pp.3-5.

Forse la modalità di pensiero dell'ES è del tipo *GPS*, il che significa soltanto che forse può, in linea di principio, essere formulata in qualche modo (non che il *GPS* sia neppure un candidato per una possibile formulazione). Forse essa converte un problema del tipo:

> Il numero di pesci di Gianni è due volte il numero dei criceti di Maria. Se Maria ha tre criceti, quanti pesci ha Gianni?

nei suoi termini peculiari, per esempio in

$$x = 2y \; ; \; y = 3$$

e lo risolve usando tecniche di elaborazione dati e di manipolazione di simboli caratteristiche del «pensiero» alla *GPS*. Ma allora non è possibile che un simile meccanismo abbia idea di cosa siano pesci e criceti, o di cosa significhi essere un ragazzo di nome Gianni ecc. Né è possibile riconvertire la rappresentazione simbolica del problema nella sua enunciazione originale. Ma la soluzione dei problemi da parte dell'uomo, forse anche quella in apparenza più meccanica e di routine, coinvolge tanto il modo di pensare destro quanto il sinistro. E, certamente, la comunicazione umana diretta coinvolge entrambi in modo cruciale.

È fin troppo facile, specialmente per gli scienziati del computer, venire ipnotizzati dal «fatto» che le espressioni linguistiche si possono rappresentare come sequenze lineari di simboli. Da questo «fatto» è facile dedurre che la comunicazione linguistica sia interamente un affare dell'emisfero sinistro. Ma la parlata umana ha anche una melodia, e il suo canto comunica tanto quanto il libretto. La musica è una provincia dell'emisfero destro, e lo stesso dicasi per il riconoscimento del valore dei gesti. Come per la comunicazione scritta, la sua funzione è certamente, almeno in gran parte, di stimolare ed eccitare particolarmente le fantasie uditive dello scrittore come del lettore.

Forse non sapremo mai se la congettura che una parte di noi pensi in termini di strutture simboliche che possono soltanto venire avvertite, ma non spiegate in modo utilizzabile, sia vera o falsa. Gli scienziati, naturalmente, hanno in orrore le ipotesi che sembrano non essere falsificabili. Eppure può darsi che, in qualche profonda accezione della parola verità, l'ipotesi sia vera. Forse essa ci aiuta a spiegare perché rimaniamo per tutta la vita stranieri per noi stessi e per gli altri, perché ogni parola del nostro lessico sia avviluppata in almeno qualche residuo di mistero, e perché ogni tentativo di risolvere i problemi della vita con strumenti interamente razionali fallisca sempre.

Ma la deduzione che desidero ricavare qui dalla mia congettura è che, siccome non possiamo sapere che essa è falsa più di quanto possiamo sapere che è vera, non abbiamo diritto alla *hubris* espressa in modo tanto altisonante dall'intellighenzia artificiale. Perfino motivi di calcolo impongono di pensare che si debba restare in soggezione, di fronte allo spettacolo misterioso che è l'uomo nella sua interezza; aggiungerei anche, a quello che è la formica nella sua interezza.

C'è stato un tempo in cui la fisica ha sognato di riuscire a spiegare l'intera realtà usando un unico formalismo esauriente. Leibniz ci insegnò che, se conoscessimo posizione e velocità di ciascuna particella elementare nell'universo, potremmo prevedere tutto il corso futuro dell'universo. Ma poi Werner Heisenberg dimostrò che gli stessi strumenti che l'uomo deve usare per poter misurare i fenomeni fisici perturbano quei fenomeni, e che è perciò impossibile, anche in linea di principio, conoscere simultaneamente l'esatta posizione e l'esatta velocità anche di una sola particella elementare. In questo modo, egli non falsificò l'ipotesi di Leibniz. Mostrò, invece, che la sua premessa maggiore non era realizzabile. Ciò, naturalmente, fu sufficiente a mandare in frantumi il sogno di Leibniz. Poco tempo dopo, Kurt Gödel mise in luce la precarietà dei fondamenti della matematica, e della logica stessa, dimostrando che ogni sistema formale interessante ha asserzioni la cui verità o falsità non può essere decisa, se si usano soltanto i mezzi formali del sistema stesso; in altre parole, che la matematica dovrà necessariamente essere sempre incompleta. Da questo e da altri risultati di Gödel consegue che «La mente umana è incapace di formalizzare (o meccanizzare) tutte le sue intuizioni matematiche. Vale a dire: se è riuscita a formalizzarne alcune, questo stesso fatto genera di per sé nuove conoscenze intuitive» [11].

Il cosiddetto «principio d'indeterminazione» di Heisenberg e il «teorema d'incompletezza» di Gödel hanno generato delle terribili onde d'urto, che hanno attraversato il mondo della fisica, della matematica e della filosofia della scienza. Ma nessuno ha smesso per questo di lavorare. I fisici, i matematici e i filosofi hanno accettato, più o meno di buon grado, la verità innegabile secondo cui esistono dei limiti al grado in cui il mondo può venir capito in termini solamente leibniziani.

Già troppe parole sono state spese sulle presunte implicazioni dei risultati di Heisenberg e Gödel per l'intelligenza artificiale. Non è mia intenzione, qui, contribuire a quel dibattito. Ma c'è un senso in cui per la psicologia e l'intelligenza artificiale può essere vantaggioso seguire l'esempio della ritrovata umiltà della matematica e della fisica moderna: esse dovrebbero riconoscere che «mentre i vincoli e le limitazioni della logica non esercitano la loro forza sulle cose del mondo, essi vincolano e limitano quelle che vogliono essere descrizioni e interpretazioni difendibili delle cose» [12]. Se lo riconoscessero, potrebbero poi affrontare il successivo passo liberatorio di riconoscere che verità non è equivalente a dimostrabilità formale.

La lezione che ho cercato di insegnare qui non è che la mente umana sia soggetta a indeterminazioni «alla Heisenberg» — anche se magari è vero —

[11] H. Wang, *From Mathematics to Philosophy* («Dalla matematica alla filosofia»), Humanities Press, New York 1974, p. 324. Lo stesso Kurt Gödel si è riferito a questo, nel dicembre 1951, come a uno dei «due più importanti risultati rigorosamente dimostrati sulla mente e la macchina». L'altro è che o esistono certi interrogativi formali cui né l'uomo né la macchina possono rispondere, oppure la mente umana può rispondere a interrogativi formali cui la macchina non può rispondere.

[12] D.C. Denett, *The Abilities of Men and Machines* («Le capacità dell'uomo e della macchina»). Lavoro presentato alla American Philosophical Association, 29 dicembre 1970.

e che perciò non potremo mai comprenderla completamente in termini di quella riduzione a fenomeni individuali che Leibniz aveva in mente. La lezione è piuttosto che quella parte della mente umana che comunica con noi in termini razionali e scientifici è essa stessa uno strumento che perturba ciò che osserva, e specialmente il suo partner privo di voce, l'inconscio, mentre media fra quest'ultimo e il nostro io conscio. I suoi vincoli e le sue limitazioni circoscrivono ciò che si può considerare descrizione e interpretazione razionale — di nuovo, se volete, scientifica — delle cose del mondo. Simili descrizioni non potranno perciò mai essere complete, non più di quanto una partitura musicale possa essere una descrizione o una interpretazione completa anche della più semplice canzone.

Ma, ed è questo il piacere salvifico di cui uno scientismo insolente e arrogante cerca di derubarci, noi arriviamo a conoscere e a capire non soltanto per mezzo dei meccanismi dell'io conscio. Noi siamo capaci di ascoltare con il terzo orecchio, di percepire la verità vivente, che è verità al di là di qualsiasi standard di dimostrabilità. Ed è *questo* tipo di comprensione, con il tipo di intelligenza che ne deriva, che io sostengo essere al di là delle capacità simulative di un computer.

Abbiamo l'abitudine — che talvolta ci torna utile — di parlare dell'uomo, della mente, dell'intelligenza e di altri concetti universali simili. Ma in modo graduale, e perfino furtivo, la nostra mente è stata infettata da quello che A.N. Whitehead chiamò «inganno della concretezza malriposta». Finiamo per credere che questi termini teorici siano in ultima analisi interpretabili come se fossero osservazioni, che nel «futuro visibile» avremo degli strumenti ingegnosi in grado di misurare gli «oggetti» cui questi termini si riferiscono. «La mente», però, non esiste; esistono solo menti individuali, ciascuna appartenente non già all'«uomo», ma a esseri umani individuali. Ho sostenuto che l'intelligenza non può venir misurata con aste millimetrate collocate lungo un *continuum* monodimensionale. L'intelligenza può venir discussa in modo utile soltanto in termini di domini di pensiero e azione. Da questo io ricavo la conclusione che non può essere utile, per dire il meno, basare un lavoro serio sulla nozione di «quanta» intelligenza sarà possibile dare a un computer. I dibattiti basati su idee come queste — per esempio: «Il computer supererà mai l'uomo in intelligenza?» — sono condannati alla sterilità.

Ho sostenuto che l'essere umano individuale, come ogni altro organismo, è definito dai problemi che affronta. L'uomo è unico in virtù del fatto che egli deve necessariamente affrontare problemi che sorgono dalle sue, uniche, necessità biologiche ed emotive. L'individuo umano è in uno stato di continuo divenire. Il mantenimento di questo stato, della sua umanità, anzi la sua sopravvivenza, dipende in modo cruciale dal suo vedersi, e dal suo essere visto dagli altri esseri umani, come un essere umano. Nessun altro organismo, e certamente nessun computer, può venir portato ad affrontare dei problemi genuinamente umani in termini umani. E dal momento che il dominio dell'intelligenza umana è determinato — a parte un piccolo insieme di problemi formali — dall'umanità dell'uomo, qualsiasi altra intelligenza, per quanto grande, deve necessariamente essere estranea al dominio umano.

Ho sostenuto che c'è un aspetto della mente umana, l'inconscio, che non può essere spiegato dai concetti primitivi dell'elaborazione di dati, i processi informatici elementari, che associamo al pensiero formale, al calcolo e alla razionalità sistematica. Eppure, questi processi siamo costretti a usarli per le spiegazioni, le descrizioni e le interpretazioni scientifiche. È indispensabile, perciò, che restiamo consapevoli della povertà delle nostre spiegazioni e della loro portata strettamente limitata. È sbagliato affermare che sia possibile un rendiconto scientifico dell'«uomo nella sua interezza». Ci sono cose che vanno al di là del potere di comprensione completa della scienza.

Il concetto di «intelligenza estranea a certi domini di pensiero e azione» è cruciale per capire quelli che sono forse i limiti più importanti dell'intelligenza artificiale. Ma quel concetto si applica alle relazioni degli uomini fra loro, tanto quanto alle macchine e al loro rapporto con l'uomo. Perché la socializzazione umana, pur essendo fondata sulla costituzione biologica comune a tutti gli uomini, è fortemente condizionata dalla cultura. E le culture dell'uomo sono molto diverse fra loro. Innumerevoli studi confermano ciò che deve essere ovvio per tutti gli osservatori del palcoscenico umano, esclusi soltanto i più miopi: «L'influenza della cultura è universale nel senso che, da certi punti di vista, un uomo impara a diventare simile a tutti gli altri; ed è particolare nel senso che un uomo allevato in una particolare società impara a diventare, sotto certi aspetti, simile a tutti gli uomini della sua società e non a quelli di altre» [13]. Gli autori di questa citazione, studiosi della società giapponese che hanno vissuto per molti anni tra i giapponesi, proseguono facendo le seguenti osservazioni:

> Nella normale vita familiare, in Giappone, c'è un'enfasi sulla mutua dipendenza e sul fare assegnamento sugli altri, mentre negli Stati Uniti l'enfasi è sull'indipendenza e sull'affermazione di sé. [...] In Giappone, il bambino è visto di più come un organismo biologico separato che fin dall'inizio, per potersi sviluppare, ha bisogno di essere inserito in rapporti sempre più interdipendenti con gli altri. Negli Stati Uniti, il bambino è visto piuttosto come un organismo biologico dipendente che, per potersi sviluppare ha bisogno di essere reso sempre più indipendente dagli altri.
>
> Il bimbo giapponese sembra passivo, e giace tranquillo con occasionali vocalizzi di insoddisfazione, mentre sua madre, nel badargli, tende più a cullarlo, a prenderlo in braccio. Sembra che cerchi di calmarlo e tranquillizzarlo, e che comunichi con il bimbo in modo fisico piuttosto che verbale. Invece il bimbo americano è più attivo, più lietamente vocioso, e deciso a esplorare il suo ambiente, e sua madre, nel badargli, tende più a guardarlo e a parlargli. Sembra che stimoli il bimbo all'attività e alla risposta vocale. È come se la madre americana volesse avere un bambino ciarliero, attivo, e come se la madre giapponese volesse avere un bambino tranquillo e soddisfatto. Dal punto di vista dello stile di comportamento delle madri nelle due culture nella cura del bambino, esse ottengono quel-

[13] W. Caudill - H. Weinstein, *Maternal Care and Infant Behavior in Japan and in America* («Cure materne e comportamento infantile in Giappone e negli Stati Uniti»), «Psychiatry» 32 (1967); pp. 12-43. Ristampato in C.S. Lavatelli - F. Stendler (a cura di), *Readings in Child Behavior and Development* («Letture sul comportamento e sullo sviluppo del bambino»), Harcourt Brace Jovanovich, New York 1972-3, p. 78.

lo che evidentemente vogliono |...| Una buona parte dell'apprendimento culturale ha avuto luogo prima di tre-quattro mesi di vita. |...| I bimbi a questo punto hanno imparato a essere bimbi giapponesi o americani, in rapporto alle attese delle madri riguardo al loro comportamento.
I giapponesi |adulti| tendono di più a formare «gruppo» e a essere interdipendenti nei rapporti con gli altri, mentre gli americani tendono di più all'«individualismo» e all'indipendenza. |...| I giapponesi sono passivi e si tengono in disparte se paragonati agli americani, che appaiono più aggressivi e desiderosi di affermarsi. |...| I giapponesi sono più sensibili a, e fanno uso conscio di, molte forme di comunicazione non verbali nei rapporti umani, attraverso i gesti e la vicinanza fisica, rispetto agli americani, che usano per lo più la comunicazione verbale in un contesto di separazione fisica.
Se questi distinti modelli di comportamento sono già in parte appresi all'età di tre-quattro mesi, e continuano a esserlo per tutta la vita della persona, allora è molto probabile che vi siano zone importanti di differenza nella risposta emotiva tra gli uomini di una cultura e quelli di un'altra. Queste differenze non possono essere soggette facilmente a un controllo conscio e, in gran parte inconsapevolmente, danno tono e colore al comportamento umano. Queste differenze |...| possono anche trasformarsi in perplessità e antagonismo, quando la gente tenta di comunicare attraverso la barriere emozionali della cultura [14].

Differenze così profonde nella prima educazione influenzano in modo cruciale le intere società interessate. E vengono, naturalmente, trasmesse da una generazione alla successiva, e così perpetuate. Esse devono necessariamente aiutare anche a stabilire che cosa i membri delle due società sappiano dei loro mondi, quali norme e valori culturali vadano considerati «universali», e quindi cosa in ciascuna cultura debba e non debba essere considerato un fatto. Esse determinano, per esempio (e questo è particolarmente significativo nel confronto tra le regole sociali giapponesi e americane) quali conflitti vadano considerati privati e quali pubblici, e quindi quali siano i criteri di giudizio appropriati alla difesa degli interessi umani, nonché quali siano questi interessi. I giapponesi per tradizione preferiscono appianare i litigi, anche quelli per cui la legge stabilisce norme precise, con sistemi che un occidentale considererebbe informali. In realtà, questi sistemi sono assai spesso regolati essi stessi da rigorosi requisiti ritualistici, che non sono codificati in modo esplicito in alcun libro, ma sono noti a qualunque giapponese che appartenga alla classe sociale interessata. Questo tipo di conoscenza si acquista col latte materno, e attraverso l'intero processo di socializzazione, che è esso stesso legato così intimamente all'apprendimento, da parte dell'individuo della sua madrelingua. Non sono cose che si possono imparare dai libri; e non si possono esplicitare, se non proprio vivendole.

Un giudice americano, perciò, non importa quanto intelligente e di ampie vedute, non potrebbe sedere in un tribunale familiare giapponese. La Corte Suprema degli Stati Uniti lo ha riconosciuto attivamente, quando ancora aveva giurisdizione su territori lontani. Per esempio, nel caso di Diaz contro Gonzales, giudicato originariamente a Portorico, la corte rifiutò l'annullamento

[14] *Ivi*, pp. 80.

della sentenza della corte della giurisdizione originaria, cioè della corte indigena. Il giudice Oliver W. Holmes, nello scrivere la motivazione del giudizio della Corte, affermò:

> Questa Corte ha riaffermato molte volte il rispetto dovuto alla comprensione dei tribunali locali su questioni di interesse puramente locale. Questo è tanto più vero quando si tratti delle decisioni di una Corte che esprime il retaggio di — ed è cresciuta in — un sistema diverso da quello che prevale qui. Quando osserviamo tale sistema dall'esterno, esso appare come una parete di pietra, ogni parte in linea con tutte le altre, tranne là dove la nostra educazione locale può portarci a vedere subordinazioni cui siamo abituati. Ma per chi viene allevato al suo interno l'enfasi mutevole, le ipotesi tacite, le consuetudini non scritte, mille influenze ricevute soltanto dalla vita possono conferire alle diverse parti valori del tutto nuovi, che la logica e la grammatica non avrebbero mai potuto estrarre dai libri [15].

Ogni intelligenza umana è quindi estranea a moltissimi domini di pensiero e azione. In ogni cultura ci sono vaste zone di interesse autenticamente umano, in cui nessun membro di un'altra cultura potrà mai prendere decisioni responsabili. Non è che l'*outsider* proprio non possa decidere — potrà sempre lanciare una moneta, per esempio — quanto piuttosto che la *base* su cui dovrebbe decidere deve per forza essere non appropriata al contesto in cui la decisione va presa.

Cosa c'è di più ovvio del fatto che, per quanta intelligenza un computer possa avere, comunque l'abbia acquisita, esso dovrà essere sempre e necessariamente estraneo a qualsiasi faccenda autenticamente umana? Il solo fatto di porre la domanda «Che cosa sa un giudice (o uno psichiatra) che non possiamo dire a un computer?» è un'oscenità mostruosa. Il fatto stesso che possa andare in stampa, anche se per lo scopo di mostrarne la natura patologica, è un sintomo della follia dei nostri tempi.

I computer possono prendere decisioni giuridiche, e formulare giudizi psichiatrici. Essi sanno lanciare la moneta in modo più sofisticato di quanto non possa anche il più paziente fra gli esseri umani. Il punto è che questi compiti *non devono* essere loro affidati. Può anche darsi che in certi casi essi sappiano arrivare a decisioni «corrette»; ma sempre e necessariamente su basi che nessun essere umano dovrebbe accettare.

Ci sono stati molti dibattiti su «Il computer e la mente». La mia conclusione qui è che le questioni pertinenti a questo tema non sono né tecnologiche né matematiche; sono etiche. E non possono essere risolte ponendo domande che incominciano con «Può...?». I limiti di applicabilità dei computer possono in definitiva esser stabiliti soltanto a base di «Si deve...?». La più elementare regola che emerge è che, visto che non abbiamo alcun modo per rendere saggi i computer, non dobbiamo dar loro compiti che richiedano saggezza.

[15] Diaz v. Gonzales, 261 U.S. 102 (1923), Per Holmes, O.W. Devo questa citazione al professor Paul Freund della *Law School* dell'Università di Harvard.

9
PROGRAMMI INCOMPRENSIBILI

Nei capitoli precedenti abbiamo discusso cosa siano i computer, da dove venga la loro potenza, e come possano venire usati per costruire modelli e incarnare teorie. Mi sono interessato dei tentativi di far comportare il computer in modo intelligente, ma non della sua applicazione a problemi numerici di carattere pratico. Ho descritto e discusso parte della ricerca di punta sulla soluzione dei problemi, sulla simulazione dei processi cognitivi e sulla comprensione del linguaggio naturale da parte dei computer, e ho descritto quelle che sono le visioni dell'intellighenzia artificiale, citandone spesso i leader riconosciuti. Non ho citato, o quasi, i fallimenti patiti dai ricercatori sull'intelligenza artificiale, perché i fallimenti che punteggiano ogni sforzo di ricerca non sono necessariamente un motivo di disperazione; al contrario, spesso essi arricchiscono il terreno, da cui in seguito spuntano idee migliori. Inoltre, il fatto che qualcosa non sia stato ancora fatto, o anche che un tentativo di farlo sia fallito, non dimostra che non lo si possa fare. Nel tentativo di evitare le discussioni banali e quelle sterili — per esempio, sulla possibilità che i computer possano «in linea di principio» imparare a svolgere questa o quell'attività — ho forse addirittura dato l'impressione di pensare che il potenziale dei computer (benché non ancora sfruttato a fondo) sia più grande di quanto, in realtà, penso che sia. Nel sostenere che ci sono problemi che gli uomini affrontano ma che le macchine non potranno mai affrontare, e che l'uomo perciò viene a conoscere cose che nessuna macchina conoscerà mai, forse ho creato l'impressione che tutti i problemi che possono presentarsi sia all'uomo che alla macchina siano potenzialmente risolvibili dalle macchine. Non è questo che intendo dire.

Le conquiste dell'intellighenzia artificiale sono per lo più dei trionfi della tecnica. Il loro contributo alla psicologia della conoscenza o alla risoluzione pratica dei problemi è stato modesto. Di certo, ci sono stati quelli che si potrebbero chiamare sottoprodotti, come dei miglioramenti nei linguaggi di programmazione di livello superiore, che hanno preso spunto da problemi di intelligenza artificiale e sono poi confluiti nel corso principale della scienza dei

computer. Ma non sono certo questi i risultati che l'intellighenzia artificiale ha continuato a promettere per il «futuro visibile» in tutti questi anni. Con poche eccezioni, non c'è stato alcun risultato, dopo oltre vent'anni di ricerca sull'*IA*, che abbia trovato applicazione nell'industria in generale, o nell'industria dei computer in particolare.

Due eccezioni sono i notevoli programmi *Dendral* e *Macsyma* esistenti alla Stanford University e al MIT, rispettivamente [1]. Entrambi questi programmi svolgono funzioni altamente tecniche, la cui discussione va ben al di là della portata di questo libro. Ma qualche parola su di essi la si può dire.

Dendral interpreta i dati di uscita degli spettrometri di massa, strumenti usati per l'analisi di molecole chimiche. È consuetudine che i giovani chimici, appena ottenuto il dottorato, vengano impiegati nella deduzione della struttura chimica di molecole introdotte in questo strumento, a partire dai cosiddetti «spettri di massa» che esso produce. Il loro problema è grosso modo simile a quello della ricostruzione della vita di un villaggio preistorico a partire dai resti recuperati dagli archeologi. C'è, tuttavia, un'importante differenza tra i due problemi: per la spettrometria di massa esiste una teoria; vale a dire, si sa come lo strumento produce i suoi dati d'uscita, se si analizza un determinato composto chimico. È possibile quindi giudicare se una soluzione proposta sia giusta o sbagliata, deducendo dalla teoria lo spettro che lo strumento produrrebbe se il composto fosse quello che la soluzione provvisoria suggerisce che sia. Purtroppo, limiti di precisione fanno sì che questo processo di valutazione non sia assolutamente esatto. Ciò nonostante, l'analista è in una situazione migliore rispetto all'archeologo, che non ha alcun metodo sicuro per verificare le sue ipotesi. In termini generali, dunque, *Dendral* è un programma che analizza spettri di massa e produce descrizioni della struttura di molecole che, con probabilità molto elevata, hanno dato origine a quegli spettri. La «competenza» del programma è pari a o maggiore di quella dei chimici, nell'analizzare certe classi di molecole organiche.

Macsyma è, secondo gli standard attuali, un programma di enormi dimensioni per fare manipolazioni matematiche simboliche. Può manipolare espressioni algebriche contenenti variabili formali, funzioni e numeri. Può differenziare, integrare, trovare limiti, risolvere equazioni, separare polinomi in fattori, sviluppare funzioni in serie di potenze, e così via. Fa tutte queste cose in forma simbolica, non numerica. Così, per esempio, dato il problema di calcolare l'integrale

[1] Su *Dendral,* vedere B. Buchanan - G. Sutherland - E.A. Feigenbaum, *Heuristic Dendral: A Program for Generating Explanatory Hypotheses in Organic Chemistry* («Il programma euristico *Dendral:* un programma che genera ipotesi di spiegazione nella chimica organica»), in B. Meltzer, a cura di, *Machine Intelligence* («L'intelligenza meccanizzata»), American Elsevier, New York 1969.
Su *Macsyma* vedere J. Moses, *Symbolic Integration, the Stormy Decade* («Integrazione simbolica, il decennio tempestoso»), in *Communications of the Associations for Computing Machinery,* vol. 14, 1971, nr. 8; W.A. Martin - R.J. Fateman, *The Macsyma System* in *Proceedings of the Second Symposium on Symbolic and Algebraic Manipulation,* Assoc. for Comp. Machinery, New York 1971.

$$\int \frac{dx}{a+bx} \ ,$$

esso produrrà

$$\frac{log\ (a+bx)}{b} \ .$$

Naturalmente, se si assegnano valori numerici a tutte le variabili esso produrrà il valore numerico dell'intera espressione, ma questo è per il programma un compito relativamente banale. Ancora una volta, i dettagli tecnici vanno al di là della portata di questa discussione. Quello che importa qui è che, proprio come per *Dendral*, esistono delle teorie ben assestate su come si debbano fare le trasformazioni richieste. Cosa importantissima, che vale particolarmente per l'integrazione simbolica, è possibile (differenziando) controllare se una soluzione proposta sia in realtà una soluzione; e, per l'integrazione, il test è assoluto. E proprio come per *Dendral*, il lavoro che *Macsyma* fa è un lavoro normalmente svolto da specialisti altamente qualificati.

Questi due programmi hanno un debito significativo nei confronti del movimento dell'intelligenza artificiale. Entrambi usano metodi euristici per la risoluzione dei problemi, in due modi distinti. In primo luogo, quando si incominciò a progettare questi programmi, le teorie su cui essi sono ora basati non erano abbastanza sviluppate da poter essere modellate in termini di procedure efficaci. Eppure, i ricercatori riuscirono nel loro compito. Un primo problema fu perciò quello di scoprire quali metodi euristici gli esperti usassero nel fare il loro lavoro. Le versioni iniziali di questi programmi erano una commistione di algoritmi che incorporavano gli aspetti ben compresi del problema, e che codificavano tutte le tecniche euristiche che si potevano mettere insieme osservando il lavoro degli esperti. Col procedere della ricerca, tuttavia, le componenti euristiche dei programmi vennero capite sempre meglio, fino a essere convertite in arricchimenti delle relative teorie. Entrambi i programmi furono così modificati gradualmente, fino a essere in pratica completamente basati su teorie. In secondo luogo, metodi euristici vennero e continuano a essere usati nei due programmi, per motivi di efficienza. Entrambi i programmi generano dei sottoproblemi che, sebbene siano risolvibili in linea di principio con semplici metodi algoritmici, si affrontano più facilmente classificandoli come suscettibili di soluzione per mezzo di qualche funzione speciale, e passandoli a quella funzione perché li risolva. Lo sviluppo e il miglioramento di entrambi gli usi di metodi euristici, così come molti dei metodi stessi, sono un prodotto della ricerca sull'intelligenza artificiale.

Questi due programmi si differenziano dalla maggior parte degli altri programmi dell'*IA* proprio per il fatto di essere saldamente fondati su profonde

teorie. I contributi principali alle fondamenta teoriche di *Dendral* sono venuti da Joshua Lederberg, genetista e premio Nobel, mentre la base teorica di *Macsyma* è principalmente opera del professor Joel Moses del MIT, un matematico di grande talento e competenza.

Ci sono, naturalmente, molte altre applicazioni importanti e ben riuscite dei computer. Computer, per esempio, controllano intere raffinerie di petrolio, guidano navicelle spaziali, sorvegliano e in larga misura controllano l'ambiente in cui gli astronauti svolgono le loro attività. I loro programmi si fondano sulla teoria matematica del controllo e su teorie fisiche saldamente fondate. Questi programmi basati su teorie godono dell'enorme vantaggio che, quando si comportano male, i loro sorveglianti umani possono riconoscere che le loro prestazioni non corrispondono correttamente ai dettami della teoria, e possono usare la teoria per diagnosticare la ragione dell'errore.

Ma quasi tutti i programmi esistenti, e specialmente quelli più grandi e più importanti, non sono basati sulla teoria in questo modo. Essi sono euristici, non necessariamente nel senso che impiegano metodi euristici al loro interno, quanto nel senso che la loro costruzione è basata su regole empiriche, stratagemmi che sembrano «funzionare» nella maggior parte dei casi previsti, e su altri meccanismi *ad hoc* che vi vengono introdotti di tanto in tanto.

Il mio stesso programma, *Eliza*, era proprio di questo tipo. Lo stesso dicasi per il sistema di Winograd per la comprensione del linguaggio e, nonostante tutte le pretese in senso contrario, per il *GPS* di Newell e Simon. Quello che è molto più importante, però, è che quasi tutti gli enormi programmi per computer usati nell'industria, negli uffici governativi e nelle università sono anch'essi di questo tipo. Questi giganteschi sistemi computerizzati di solito sono stati messi insieme (non sempre si può usare la parola «progettati») da gruppi di programmatori, il cui lavoro spesso si estende su un periodo di molti anni. Nel momento in cui questi sistemi entrano in funzione, la maggior parte dei programmatori originari se ne è andata, o si è dedicata ad altri progetti. È proprio quando questi sistemi cominciano a essere utilizzati che la loro struttura interna non è più comprensibile ad alcun singolo individuo o piccolo gruppo di persone.

Norbert Wiener, il padre della cibernetica, predisse questo fenomeno in un articolo di notevole preveggenza pubblicato più di quindici anni fa. In esso scriveva:

> È perfettamente possibile che non si possa, in linea di principio, costruire alcuna macchina di cui non riusciremmo, prima o poi, a capire il comportamento in tutti i suoi aspetti. Questo però non significa affatto che riusciremo a capire quegli aspetti in un tempo sostanzialmente inferiore a quello richiesto per il funzionamento della macchina, o addirittura in un qualsiasi dato numero di anni o di generazioni.
> Può accadere che una comprensione intelligente delle modalità di funzionamento (di una macchina) avvenga soltanto molto tempo dopo il completamento del

lavoro che (le si) è affidato. |...| Questo significa che, benché le macchine siano in teoria soggette alla critica umana, questa critica può restare inefficace fino a molto tempo dopo il momento in cui sarebbe stata pertinente [2].

Quella che Norbert Wiener descriveva come una possibilità è diventata già da molto tempo una realtà. Sembra che la ragione di questo sia quasi impossibile da capire o accettare per l'uomo della strada. La sua errata concezione di cosa siano i computer, di cosa facciano e di come lo facciano è in parte imputabile alla capacità della metafora meccanicistica di permeare e penetrare in profondità l'inconscio della nostra intera cultura. Questo è un retaggio dell'impatto che hanno avuto sull'immaginazione le macchine relativamente semplici che hanno trasformato la vita durante il diciottesimo e il diciannovesimo secolo. L'idea che capire qualcosa equivalesse a capirla in termini meccanicistici divenne praticamente una «seconda natura» per tutti coloro che vivevano nei paesi industrializzati. Perfino alcuni grandi scienziati del tardo diciannovesimo secolo appoggiarono questo punto di vista. Lord Kelvin (1824-1907) scrisse: «Non sono mai soddisfatto fino a quando non riesco a fare un modello meccanico delle cose. Se riesco a farne un modello meccanico, riesco a capirle. Se non riesco ad arrivare fino in fondo con un modello meccanico, non le capisco» [3]. Un'espressione del sentimento moderno corrispondente è la credenza di Minsky, secondo cui capire la musica e le «immagini altamente significative» equivale a essere capaci di scrivere programmi per computer in grado di generarle. Ma mentre Minsky capisce benissimo che i computer non sono macchine da mettere sullo stesso piano dei meccanismi noti a Kelvin, l'uomo della strada capisce esattamente il contrario. Per lui, computer e programmi sono «meccanici» nello stesso semplice senso in cui lo sono i motori a vapore o gli alberi di trasmissione delle automobili.

Questa credenza — praticamente universale tra i profani — è rinforzata dallo slogan, spesso ripetuto dagli stessi scienziati del computer, secondo cui «se un procedimento non è formulato con precisione assoluta, non lo si può far eseguire a un computer». Questo slogan, però, è vero soltanto in un'interpretazione molto rigorosa — e assai insolita — di cosa significhi «formulare un procedimento». Se uno, per esempio, lanciasse una serie di bit a caso nella memoria di un computer, e la facesse interpretare dal computer come un programma, allora, supponendo che «funzionasse», quella serie di bit sarebbe una «formulazione» di qualche procedimento. Ma nella conversazione quotidiana, per formulazione di un programma si intende che un qualche soggetto (presumibilmente umano) abbia organizzato quello che deve diventare un programma, prima di affidarlo al computer. Il profano, avendo sentito lo slogan in questione, crede che il fatto stesso che un programma giri su un computer ga-

[2] N. Wiener, *Some Moral and Technical Consequences of Automation* («Alcune conseguenze morali e tecniche dell'automazione»), «Science», vol. 131, 1960, n. 3410 p. 1355.

[3] W.T. Kelvin, cit. in P.W. Bridgeman, *The Logic of Modern Physics* («La logica della fisica moderna»), Macmillan, New York 1946, p. 45.

rantisca che qualche programmatore abbia formulato, e capisca, ogni dettaglio del procedimento che quel programma incarna.

Ma questa credenza è contraddetta dai fatti. Un grande programma è, per usare un'analogia cara anche a Minsky, una rete di tribunali — cioè, subroutine — collegati tra loro in modo intricatissimo, a ciascuno dei quali vengono trasmesse prove da parte di altre subroutine. Questi tribunali vagliano i dati ricevuti, per poi trasmettere la loro sentenza ad altri tribunali ancora. I verdetti emessi da questi tribunali possono implicare, e anzi lo fanno spesso, delle decisioni su quale tribunale abbia «giurisdizione» sui risultati intermedi che sono in fase di manipolazione. Il programmatore, in questo modo, non può nemmeno sapere quale sia l'itinerario delle decisioni prese all'interno del suo stesso programma, per non parlare dei risultati intermedi o finali che esso produrrà. La formulazione di un programma è così più simile alla creazione di una burocrazia che alla costruzione di una macchina del tipo che Lord Kelvin poteva avere in mente. Nelle parole di Minsky:

> Il programmatore stesso stabilisce |...| principi «legali» che ammettono |...| «appello», ma può avere soltanto una visione molto incompleta di quando e dove, nel corso dell'esecuzione del programma, queste procedure si chiameranno fra loro. E per un particolare «tribunale», egli ha soltanto una vaga idea di alcune delle circostanze che lo faranno intervenire. In breve, una volta superato il livello di principianti, |...| i programmatori scrivono non già «sequenze» (di istruzioni), bensì regole di comportamento per gli individui di piccole società. Per quanto ci provi, il programmatore sarà spesso incapace di prevedere in anticipo tutti i dettagli delle loro interazioni. Perché è proprio per questo, in fin dei conti, che ha bisogno del computer [4].

Minsky prosegue facendo le seguenti osservazioni, di enorme importanza:

> Quando un programma cresce in potenza attraverso un'evoluzione di rattoppi e aggiustamenti solo parzialmente compresi, il programmatore incomincia a perdere traccia dei dettagli interni, perde la capacità di prevedere quello che accadrà, comincia a *sperare* invece di *sapere,* e osserva i risultati come se il programma fosse un individuo dotato di diversi possibili comportamenti.
> Questo è già vero oggi per alcuni grossi programmi |...| e presto il problema diventerà molto più grave |...| Programmi euristici di grandi dimensioni verranno sviluppati e modificati da parecchi programmatori, ciascuno dei quali li metterà alla prova su casi esemplari diversi, comunicando con il computer da terminali |decentrati| e inserendo informazioni in modo indipendente. Il programma aumenterà di efficacia, ma nessuno dei programmatori sarà in grado di capirlo tutto (naturalmente, non sempre questo metodo avrà successo: le interazioni potrebbero renderlo peggiore, nel qual caso nessuno saprebbe più rimetterlo in se-

[4] M. Minsky, *Why Programming is a Good Medium for Expressing Poorly Understood and Sloppily Formulated Ideas* («Perché la programmazione è un buon sistema per esprimere idee mal comprese e formulate in modo impreciso») in M. Kramper e P. Seeitz (a cura di), *Design and Planning II* («Progetto e programmazione II»), Hasting House, New York 1967, p. 120.

sto!). Adesso vediamo qual è il problema di affermazioni come «il computer fa soltanto ciò che il suo programmatore gli ha detto di fare»: «il programmatore» non esiste [5].

Noi non riusciamo a capire, per ritornare a un punto precedente, come un programma del tipo che Minsky descrive qui — un programma, mettiamo, che compone della «grande» musica — ci possa aiutare a «capire» la musica, quando già il programma va al di là della nostra comprensione.

Ma, cosa più importante, se il programma ha superato la comprensione di coloro che l'hanno creato, che cosa può significare dire che «aumenta di efficacia» oppure che «è reso peggiore?» Come insegnanti (e siamo tutti insegnanti, in un modo o in un altro) noi naturalmente speriamo sempre che i nostri allievi aumenteranno di efficacia, nel loro modo di trattare gli argomenti oggetto del nostro tutoraggio. E di solito non pretendiamo di «capire» i procedimenti di cui vogliamo incoraggiare la crescita nei nostri studenti, vale a dire, di capirli bene quanto capiamo, per dire, il funzionamento di un orologio. Inoltre, investiamo davvero le nostre speranze nei nostri studenti, facciamo affidamento su di loro e ci fidiamo di loro.

È indubbiamente questo il tipo di fiducia che Minsky ci invita a investire in complessi programmi di *IA* che aumentano di efficacia, anche se finiscono per andare al di là della nostra comprensione. Il suo consiglio è del tutto ragionevole, se si riferisce a programmi per cui abbiamo misure di prestazione che ci permettano di dire, e di dire in un tempo sufficientemente breve, quando questi programmi stanno operando con un comportamento non accettabile o quando, qualsiasi sia la ragione, essi non meritano più la nostra fiducia. Come abbiamo osservato in precedenza, i programmi che sono sostanzialmente modelli di teorie ben assodate cadono in questa categoria, anche se, come può accadere, non c'è alcun gruppo di programmatori che abbia una comprensione dettagliata delle caratteristiche interne dei programmi stessi. Lo stesso vale per quei programmi le cui disfunzioni possono essere rivelate dall'osservazione del loro comportamento istante per istante; a patto, naturalmente, che vi sia in primo luogo qualcuno responsabile che sta guardando, e in secondo luogo che questo qualcuno possa intervenire in tempo per evitare disastri. Ma esistono oggi molti programmi importanti che sono molto estesi e complessi, e non soddisfano queste condizioni.

Il resoconto di Minsky, che è del tutto corretto, è quindi della massima importanza. Esso ci dice che la situazione che Norbert Wiener descriveva nel 1960 come una possibilità è diventata rapidamente, ed è ora, realtà. Le parole di Minsky, per di più, assumono un'importanza particolare perché sono state scritte da uno dei principali artefici e portavoce dell'intelligenza artificiale, ed erano volte a correggere le idee confuse degli umanisti esaltando la potenza dei computer, non i loro limiti.

[5] *Ivi*, p. 121.

Il crescente affidamento che la nostra società fa sui computer, che erano inizialmente destinati ad «aiutare» la gente a fare analisi e a prendere decisioni, ma che da ormai molto tempo hanno da un lato superato la comprensione di chi li usa, diventando d'altro canto indispensabili, è un fenomeno molto grave. Ha due conseguenze importanti. In primo luogo, si prendono decisioni con l'aiuto di (anzi, a volte lasciandole interamente a) computer che nessuno più conosce o capisce. Quindi, nessuno può sapere i criteri e le regole su cui quelle decisioni sono basate. In secondo luogo, i sistemi di regole e criteri incarnati da questi sistemi diventano immuni di fronte alle modifiche poiché, in assenza di una comprensione dettagliata del funzionamento interno di un sistema di computer, qualunque sua modifica rischia con tutta probabilità di non fare più funzionare l'intero sistema, e magari in modo irreversibile. Sistemi di questo tipo possono perciò soltanto crescere. E la loro crescita, insieme al crescente affidamento che si fa su di loro, è allora accompagnata da una crescente legittimazione della loro «base di conoscenze».

Il professor Philip Morrison del MIT ha scritto una parabola pungente su questo tema:

> Sulla parete del mio ufficio c'è una carta del mondo, stampata con il computer e quindi non bella come potrebbe farla un disegnatore. Su di essa si riconoscono chiaramente i contorni, costituiti da otto-diecimila punti, delle enormi zolle che formano la crosta della Terra e che, quando si separano, o si toccano, o scorrono l'una sull'altra, generano quasi tutti i terremoti di una certa importanza. La carta incarna quest'ultima nozione, perché i contorni punteggiati delle zolle sono costituiti da migliaia di epicentri di terremoti.
>
> L'aspetto curioso è questo: i sismologi responsabili della carta dicono, un po' come per scusarsi, che siccome i dati raccolti da loro sui terremoti erano espressi in un particolare formato standard, che rendeva più facile comunicarli alla macchina in modo da individuare i punti sulla carta, essi erano in grado di usare soltanto i loro dati. Sanno bene, è chiaro, che la sismologia è ben più vecchia di un decennio, ma sostengono che lo sforzo di collegare i dati del passato a un sistema di coordinate standard, e di tradurre in forma leggibile dal computer la vasta e variegata letteratura del periodo 1840-1961, andava al di là delle loro possibilità. Perciò essi hanno trascurato qualsiasi riferimento alla scienza anteriore al 1961, e hanno usato soltanto quei terremoti rilevati nel periodo 1961-1967 dalla rete mondiale di sismografi da loro organizzata. Il numero di questi ultimi, però, era più o meno pari a quello di tutti i terremoti rivelati prima del 1961. Hanno quindi perduto un fattore due, che non è molto dal punto di vista statistico, con il vantaggio di non dover leggere e interpretare tutte quelle oscure riviste tedesche.
>
> Questa è una parabola sul computer. Come tutte le parabole, ha una tensione interna: offre qualcosa tanto ai nemici quanto agli amici del computer. Per gli amici, è chiaro che questa superba collezione di epicentri che delineano le zolle tettoniche è probabilmente l'unico e più grande risultato di un simile studio sinottico. Per un profano, è affascinante vedere i contorni delle spaccature e delle giunzioni. Finalmente riusciamo a capire qualcosa sulla Terra nel suo complesso. Allo stesso tempo, una messa al bando così sprezzante dell'intera storia di una scienza lascia senza fiato.

La lezione è piuttosto evidente: nessuno, neppure il più maniaco propugnatore dell'elaborazione di dati con il computer, affermerebbe che tutto sia cominciato nel 1961, anche se le osservazioni moderne e compatibili col computer sono iniziate allora. Il passato è stato un prologo indispensabile; ha visto la formazione dei concetti, lo sviluppo delle tecniche, l'introduzione degli strumenti, l'idea della registrazione sistematica, e così via. Tutto questo ha spianato la strada, senza la quale sono sicuro che la «*Coast and Geodetic Survey*» e i suoi amici non sarebbero stati capaci di produrre una mappa così bella [6].

Il computer ha così cominciato a essere uno strumento per la distruzione della storia. Perché quando la società legittima soltanto quei «dati» che sono «in un formato standard» e che «sono facili da comunicare alla macchina, allora la storia, la memoria stessa, viene annullata. Il «New York Times» ha già incominciato a costruire una «banca-dati» di avvenimenti attuali. Naturalmente, soltanto quei dati che possono essere ricavati facilmente come sottoprodotti delle macchine compositrici possono venire immessi nel sistema. Col crescere del numero di abbonati a questo sistema, che imparano sempre più a fare affidamento su «tutte le notizie (che sono state un giorno) degne di stampa» [come il «Times» orgogliosamente descrive la sua politica editoriale] quanto tempo ci vorrà prima che sia il sistema a determinare che cosa va considerato «un fatto»? Prima che ogni altra conoscenza, ogni memoria venga semplicemente dichiarata illegittima? Presto verrà costruito un supersistema, basato sulla banca-dati del «New York Times» (o su una molto simile) da cui gli «storici» ricaveranno le loro deduzioni su cosa è «davvero» successo, su chi è collegato a chi, e sulla «vera» logica degli avvenimenti. Oggi c'è molta gente che non vede in questo niente di male.

Non è necessario guardare al futuro, per riempire di contenuto la parabola di Morrison. Nella recente guerra degli Stati Uniti contro il Vietnam, erano dei computer, fatti funzionare da ufficiali che non avevano la più lontana idea di cosa succedesse dentro le loro macchine, a scegliere quali villaggi dovessero essere bombardati, e quali regioni avessero una densità di Viet Cong sufficiente a farle «legittimamente» dichiarare zone di «fuoco a volontà», cioè vaste regioni geografiche in cui i piloti avevano il «diritto» di uccidere qualsiasi oggetto vivente. Naturalmente, in queste macchine potevano essere introdotti soltanto dati «leggibili dalla macchina», cioè per lo più informazioni sui bersagli provenienti da altri computer. E quando il Presidente degli Stati Uniti decise di bombardare la Cambogia e di tenere nascosta al Congresso quella decisione, i computer del Pentagono vennero «aggiustati» in modo da trasformare i rapporti veri sull'attacco, provenienti dal campo di battaglia, in quei rapporti falsi cui veniva dato accesso ai leader governativi. Il «Ministero della Verità» di George Orwell era stato meccanizzato. La storia non venne solamente distrutta, venne ricreata. E i leader governativi, che si sentivano privilegiati per il fatto di poter leggere i rapporti che uscivano dai computer del Pentago-

[6] P. Morrison, *The Mind of the Machine* («La mente della macchina»), «Technology Review» (MIT), gennaio 1973, p. 13.

no, naturalmente ci credevano. Dopo tutto, era stato niente di meno che il computer a parlare. Essi non si resero conto di essere diventati «schiavi» del computer, per usare le parole dell'ammiraglio Moorer, fino a quando quelle stesse bugie che avevano insegnato al computer a dire non ingannarono anche loro [7].

Nella guerra moderna è normale che il soldato, per esempio il pilota di un bombardiere, operi a una distanza psicologica enorme dalle sue vittime. Egli non è responsabile delle ustioni ai bambini, perché non vede mai i loro villaggi, né l'esplosione delle sue bombe, né i bambini stessi fra le fiamme. Le moderne razionalizzazioni tecnologiche della guerra, la diplomazia, la politica e il commercio (come nel caso dei videogiochi) hanno un effetto ancora più insidioso sulle scelte politiche. Non soltanto i politici hanno abdicato alla loro responsabilità di prendere decisioni, lasciandola a una tecnologia che non capiscono — sebbene conservino sempre l'illusione di essere loro, i politici, a formulare le domande sulla politica e a trovare le risposte —, ma la responsabilità è svanita proprio del tutto. Non soltanto il più anziano ammiraglio della Marina degli Stati Uniti può rendersi conto, in uno dei rari momenti di illuminazione, del fatto di essere diventato «uno schiavo di questi maledetti computer», e di non poter fare a meno di basare i suoi giudizi su «quello che dice il computer», ma per di più nessun essere umano è responsabile di ciò che esce dal computer. Gli enormi sistemi computerizzati del Pentagono, e le loro controparti altrove, nella nostra cultura, non hanno, in un senso molto pratico, alcun autore. E così non ammettono nessuna questione di giusto e di sbagliato, di legittimità o di teorie con cui si può essere d'accordo o meno. Non offrono alcuna base su cui sfidare «quello che dice la macchina». Mio padre di solito, per invocare l'autorità ultima, mi diceva «Sta scritto!». Ma allora, io leggo ciò che sta scritto, mi immagino che dietro vi sia un autore umano, ne deduco i suoi valori e infine mi trovo o meno d'accordo con lui. I sistemi di computer non tollerano esercizi di immaginazione che possano alla fine portare a un autentico giudizio umano.

Non c'è da stupirsi che uomini che vivono dalla mattina alla sera con macchine di cui sentono di essere diventati schiavi comincino a credere che gli uomini siano macchine; e così, come disse una volta un importante scienziato:

> È possibile vedere l'Uomo stesso come il risultato di [...] un processo evolutivo di robot capaci di svilupparsi, procreati da robot più semplici, e così a ritroso fino al brodo primordiale; [...] la sua condotta etica [è] qualcosa che va interpre-

[7] Secondo un reportage di Seymour Hersch nel «New York Times» del 10 agosto 1973, l'ammiraglio Thomas Moorer, capo dello staff della Presidenza, spiegò alla Commissione per i servizi armati del Senato degli Stati Uniti che gli attacchi aerei contro la Cambogia erano stati introdotti nel «grande elaboratore di dati del Pentagono» come attacchi contro il Vietnam del Sud. Il «Times», in un'altra parte dello stesso articolo, presenta una fotocopia di un rapporto sull'attacco che porta la dicitura: «Tutte le missioni indirizzate contro la Cambogia verranno programmate come contro bersagli alternativi nel Vietnam del Sud». Si riferisce anche che l'ammiraglio Moorer abbia detto alla Commissione senatoriale: «È increscioso che si sia dovuti diventare schiavi di questi maledetti computer».

tato in termini dell'azione dei circuiti |...| dell'Uomo nel suo ambiente: una macchina di Turing con due sole retroazioni ben determinate, un desiderio di giocare e un desiderio di vincere [8].

Ci si aspetterebbe che un gran numero di individui, vivendo in una società in cui forze anonime, e quindi irresponsabili, formulano le grandi domande del giorno e circoscrivono il campo delle possibili risposte, provasse una sorta di senso di impotenza e cadesse vittima di un furore irrazionale. E questa aspettativa certamente la vediamo realizzata tutt'intorno a noi, nei *campus* universitari e nelle fabbriche, nelle case e negli uffici. Le sue manifestazioni sono il sabotaggio da parte dei lavoratori del prodotto della loro fatica, l'irrequietezza e la mancanza di scopo tra gli studenti, la criminalità sulla strada, la fuga in mondi onirici indotti dalle droghe, e così via. Eppure c'è anche una reazione alternativa molto diffusa; da un certo punto di vista, sembra rassegnazione, ma da un altro è quella che molto tempo fa Erich Fromm chiamò «fuga dalla libertà».

Il «bravo tedesco» dei tempi di Hitler poteva dormire più tranquillo perché «non sapeva» di Dachau. Non lo sapeva, ci ha detto in seguito, perché il sistema nazista era così ben organizzato da tenerglielo nascosto. (Curiosamente, però, io che ero adolescente in quella stessa Germania sapevo di Dachau. Io pensavo di avere ragione di temerla). Naturalmente, la vera ragione per cui il bravo tedesco non sapeva è che egli non aveva mai sentito come un suo dovere quello di chiedere cosa fosse successo al suo vicino ebreo il cui appartamento si era improvvisamente liberato. Il professore universitario che vide improvvisamente realizzato il suo sogno di essere promosso al rango di ordinario, non chiese come mai la sua preziosa poltrona fosse tutt'a un tratto diventata vacante. Alla fine, tutti i tedeschi diventarono delle vittime di quello che era loro «accaduto».

Oggi, anche i manager che rivestono i ruoli di più alto livello si presentano come vittime innocenti di una tecnologia per cui non accettano alcuna responsabilità e che non fanno neanche finta di capire (ci si deve proprio chiedere, comunque, come mai non sia mai capitato all'ammiraglio Moorer di domandare che effetto avessero i milioni di tonnellate di bombe che, a quanto diceva il computer, si stavano sganciando sul Vietnam). Il Segretario di Stato americano, dottor Henry Kissinger, nello spiegare che egli non avrebbe potuto sapere degli «orrori della Casa Bianca» rivelati dall'indagine Watergate, si rammaricò «della orribilità degli avvenimenti e della tragedia che ha colpito tante persone».

La tragedia così descritta ebbe un'azione, ma nessun attore. Soltanto gli «avvenimenti» sono stati «orribili», non gli individui o gli ufficiali. In questo scenario

[8] Non darò il riferimento per questa citazione. Il brano venne scritto nel 1952. Quando ho conosciuto l'autore, molti anni dopo, egli era un uomo anziano, gentile e umano. Non posso credere che egli fosse disposto, negli ultimi anni della sua vita, a sottoscrivere ancora l'opinione qui citata.

privo di vita, la parodia della legge e l'inganno delle persone non erano stati provati e decisi a tavolino; erano semplicemente «accaduti» [9].

Il mito dell'ineluttabilità tecnologica, politica e sociale è un potente tranquillante per la coscienza. La sua funzione è quella di togliere la responsabilità dalle spalle di chiunque ci creda davvero.

Ma nella realtà ci *sono* attori!

Per esempio, una circolare programmatica distribuita al corpo insegnante e al personale dal direttore di un importante laboratorio di informatica di un'importante università diceva quanto segue [10].

> La maggior parte delle nostre ricerche è stata sostenuta, e probabilmente continuerà a esserlo, dal Governo degli Stati Uniti, e dal Dipartimento della Difesa in particolare. Il Dipartimento della Difesa, come altre agenzie del nostro governo, è impegnato nello sviluppo e nella gestione di sistemi complessi che hanno un potenziale distruttivo molto grande e che, sempre più, vengono comandati e controllati usando computer digitali. *Questi sistemi sono responsabili*, in larga misura, del mantenimento di quella pace e stabilità che esiste nel mondo, e allo stesso tempo sono capaci di dare via libera a distruzioni su una scala che è quasi impossibile per un uomo immaginare.

Notate che sono responsabili i sistemi, non le persone. Comunque, ora che è stato fatto il doveroso accenno al potenziale distruttivo, passiamo alle cose davvero importanti:

> Il ruolo cruciale dei computer può essere riconosciuto in modo più lampante nelle applicazioni militari che nei settori non militari della società, ma i più tra noi hanno riflettuto sulla dipendenza sempre crescente dai computer nel commercio e nell'industria quanto basta per proiettare un quadro in cui il funzionamento stesso della società dipende da un'esecuzione ordinata e sensata di miliardi di istruzioni elettroniche al secondo |...| Ci saranno sistemi su larga scala con memorie ad accesso casuale di milioni di parole, capaci di eseguire decine di milioni di istruzioni al secondo, in organizzazioni di ogni tipo. Per la maggior parte, questi computer saranno collegati fra loro in complessi di reti comunicative attraverso le quali avranno accesso (un accesso regolato da meccanismi di controllo derivati da ciò che alcuni di noi stanno facendo oggi) a tutte le informazioni esistenti su tutto e tutti. *E non si può opporre resistenza a questa tendenza nello sviluppo dei computer.* |...|
> Essendo padroni della programmazione e del controllo dei computer, noi in particolare potremmo svolgere un ruolo cruciale. Può benissimo darsi che nessun'altra organizzazione sia in grado di farlo come noi, eppure oggi non può esservi ruolo più importante nella scienza e nella tecnologia.

[9] J.E. Hughes, *The Wide World of Watergate* («Il vasto mondo del Watergate»), in «Newsweek», 20 agosto 1973. p. 13.

[10] Questi estratti sono citati con il permesso dell'autore. I corsivi li ho aggiunti io.

> L'importanza del ruolo deriva, come è stato osservato, dal fatto che *il computer si è incorporato* finora, e certamente continuerà a incorporarsi, nella maggior parte delle funzioni fondamentali per il sostegno, la protezione e lo sviluppo della nostra società. Già ora *non c'è possibilità di ritorno indietro*, e in pochi anni sarà chiaro che dipendiamo dall'elaborazione di informazioni fatta dai nostri computer, tanto quanto dalla crescita del grano nei campi e dall'estrazione del petrolio dai pozzi.

Non c'è neanche il più lontano accenno a una domanda per vedere se lo vogliamo, questo futuro. Sta semplicemente arrivando. Siamo impotenti, di fronte a una marea cui non si potrà, per nessuna ragione, opporre resistenza. Non è possibile tornare indietro. Non vale neppure la pena di discutere la domanda.

> Ci sono molte sfaccettature del problema generale di padronanza e controllo. Esse vanno da questioni essenzialmente filosofiche, che riguardano significati e intenzioni, e l'instaurazione di una conformità tra un progetto e il comportamento pratico di un sistema complesso, al problema quasi puramente tecnico di trovare i *bug* in una subroutine.

(nel gergo informatico, un *bug* è un errore di programmazione) [11]. Notare la visione provinciale, cioè tecnologica, della filosofia messa in mostra qui. Ma andiamo avanti:

> Non si dovrebbe essere soddisfatti di metodi di programmazione che permettono ai bug di introdursi nei programmi. È probabile che il modo migliore per eliminare i bug sia ideare metodi di programmazione intrinsecamente privi di bug. Ciò nonostante, il *debugging* dovrebbe essere un punto focale dello sforzo di ricerca intrapreso per padroneggiare la programmazione. La ragione è che la ricerca sul debugging produrrà una nuova comprensione di molti problemi di formulazione ed espressione degli intenti umani. Non si tratta solo della codificazione di una formulazione, come programma, di problemi e soluzioni. Col raffinarsi della padronanza della programmazione, ci sarà un continuo ampliamento della portata della risoluzione di problemi, un allargamento dell'universo di discorso. L'obiettivo dovrebbe essere un metodo di programmazione che sia libero da bug e tranelli a questi livelli superiori tanto quanto lo è a quelli inferiori.

Quello che è davvero notevole è che il principale — anzi, l'unico — impedimento alla «risoluzione dei problemi», perfino a «questi livelli superiori», è visto unicamente come una questione di errori tecnici. Non ci sono veri conflitti nella società. Una volta che abbiamo capito gli «intenti umani », il che è pure un problema tecnico, tutto il resto è tecnica.

[11] *Bug* in inglese significa «bacherozzo»; eliminare gli errori da un programma si dice *debugging*, (NdT).

> Quasi certamente, i computer sono più liberi da errori, nel fare qualunque cosa facciano, rispetto alle persone. Se riusciamo a creare un programma per scrivere programmi, perciò, dovremmo essere molto avanti nella strada per arrivare al *software* senza bug. È importante, però, non tacere i problemi che sorgono all'interfaccia tra l'asserzione umana di un problema e la comprensione di quel problema da parte del computer. È improbabile che il computer possa preparare un programma adeguato, a meno che la sua comprensione del problema sia completamente corretta.|...|
> Un approccio forse importante alla padronanza della programmazione e al debugging è basato sul fare modelli. |...| Al principio, il programmatore umano si aiuta con i modelli che ha a disposizione. Verso la fine, i modelli, combinati in un solo modello esauriente, fanno la maggior parte del lavoro di programmazione e di debugging, ma il programmatore umano è ancora in gioco come supervisore, o per trarre d'impaccio in situazioni difficili, per fornire una guida euristica, o chissà che. Infine, se lo sforzo ha successo, il modello diventa programmatore automatico. |...|
> Soltanto un anno o due fa, era necessario mettere tra virgolette la parola «conoscenza» quando la si usava in un contesto come questo |...| ma (all'interno di una cerchia piuttosto ristretta di scienziati del computer) oggi c'è accordo sul fatto che abbiamo raggiunto la soglia oltre la quale si può pensare che i computer abbiano conoscenze, e che le usino efficacemente e sensatamente in modi analoghi — e un giorno probabilmente saranno superiori — a quelli in cui gli esseri umani usano le loro conoscenze.|...|
> Conversando con membri del |nostro| corpo insegnante si ha una forte impressione di convergenza in una direzione nuova. Molti sembrano avvertire un comune insieme di priorità. Tradurre in parole queste sensazioni ha richiesto una buona dose di pensiero lucido, di riflessione e molte ampie discussioni.
> La convergenza delle indicazioni |...| coinvolge l'idea di rendere i computer non solo facili da usare ma anche, come si è sottolineato qui, *degni di fiducia.* |...| |Le nostre| risorse uniche saranno meglio utilizzate se dedicate totalmente ad assicurare che il futuro dominato dai computer sia quello che noi desideriamo. Forse |noi| *soli* |siamo| nella condizione di raggiungere questo importante obiettivo.

L'autore di questo documento — che, per inciso, non è una delle persone citate altrove in questo libro — sta soltanto proponendo l'esecuzione del programma tanto spesso strombazzato da altri portavoce dell'ottimismo tecnologico. Quanto scrive è, per esempio, del tutto coerente con la previsione fatta da H.A. Simon nel 1960, secondo cui:

> In un futuro molto prossimo — molto meno di venticinque anni — avremo la possibilità tecnica di sostituire qualsiasi funzione svolta dall'uomo, nelle diverse organizzazioni, con una macchina. Nello stesso tempo, avremo acquisito una teoria estesa e provata empiricamente dei processi cognitivi umani e della loro interazione con le emozioni, gli atteggiamenti e i valori umani [12].

[12] H.A. Simon, *The Shape of Automation* («La forma dell'automazione»), 1960, ristampato in *Perspectives on the Computer Revolution* («Punti di vista sulla rivoluzione del computer»), a cura di Z.W. Pylyshyn (Prentice-Hall, Englewood Cliffs, N.J., 1970).
Nello stesso lavoro, Simon predice: «La possibilità di raddoppiare le capacità cerebrali di soluzione di problemi e trattamento di informazioni non è lontana; sarebbe sorprendente se non venisse

E l'«ottimismo» messo in mostra qui non si riferisce soltanto ai computer. Il professor B.F. Skinner, leader del comportamentismo in psicologia, e di cui si dice spesso che sia lo psicologo più influente esistente oggi al mondo, ha scritto di recente:

> I risultati disastrosi del buon senso comune nella gestione del comportamento umano sono evidenti in ogni percorso di vita, dagli affari internazionali alla cura di un bambino, e continueremo a essere inetti in tutti questi campi, fino a quando un'analisi scientifica non renderà chiari i vantaggi di una tecnologia più efficace.
> Dal punto di vista comportamentistico, l'uomo può ora controllare il suo destino, perché sa cosa si deve fare e come farlo [13].

L'ultima frase non può significare altro che «Io, B.F. Skinner, so cosa si deve fare e come farlo», così come l'ultima frase del documento programmatico citato prima non può significare altro se non che «noi», che abbiamo il compito di sfruttare le nostre risorse per assicurare che il futuro dominato dal computer sia quello che «noi» desideriamo, siamo membri della cerchia piuttosto ristretta di scienziati del computer al cui interno «noi» possiamo parlare in modo schietto, e senza usare virgolette eufemistiche. È significativo che in entrambi i casi le frasi riportate siano quelle di chiusura del documento. Esse contengono, ovviamente, il messaggio finale e più importante.

Ma i messia tecnologici che, trovando impossibile fidarsi della mente umana, si sentono obbligati a costruire computer «degni di fiducia» che comprenderanno le intenzioni umane e sapranno risolvere i problemi umani, hanno anche dei concorrenti di altri quartieri. Uno dei più insigni tra questi è il professor J.W. Forrester del MIT, il padre intellettuale del movimento sistema-dinamico. In una deposizione davanti a una commissione del Congresso degli Stati Uniti [14], egli ha detto:

> La mia tesi fondamentale è che la mente umana non è adatta a interpretare il comportamento dei sistemi sociali.|...| Fino a tempi recenti non c'è stato modo di valutare il comportamento dei sistemi sociali se non con meditazioni, discussioni, dispute e congetture.

realizzata entro il prossimo decennio». Beh, è passato più di un decennio, e il cervello è rimasto misterioso come sempre. Dobbiamo immaginare che il professor Simon sia sorpreso. Talvolta, nel guardare per esempio una scena particolarmente orripilante di un film, si può riuscire a non farsi sopraffare pensando consciamente che le persone sulla scena, in fin dei conti, stanno solo recitando. Quella tecnica non funziona, in questo caso. Il professor Simon è oggi uno dei più influenti statisti della scienza in America. Quello che egli dice conta davvero.

[13] B.P. Skinner, *About Behaviorism* («Sul comportamentismo»), Alfred A. Knopf, New York 1974, pp. 234, 251.

[14] Le seguenti citazioni di Forrester sono tratte da *Testimony before the Subcommittee on Urban Growth of the Committee on Banking and Currency of the United States House of Rapresentatives, given in Washington, D.C. Oct. 7, 1970.* («Testimonianza davanti al Sottocomitato sulla crescita urbana del Comitato affari bancari e volontari della Camera dei rappresentanti degli Stati Uniti, resa a Washington, D.C., il 7 ottobre 1970»), 91esimo Congresso, seconda sessione, parte III, pp. 205-206. La si può ottenere dall'Ufficio stampa del Governo degli Stati Uniti.

In altre parole, i modi in cui Platone, Spinoza, Hume, Gandhi e tanti altri hanno ragionato sui sistemi sociali sono palesemente inferiori al modo introdotto dall'analisi dei sistemi. Il guaio è che quei modi di pensare sono basati su modelli mentali. E:

> Il modello mentale è confuso. È incompleto. È specificato in modo impreciso. Inoltre, nello stesso individuo un modello mentale cambia nel tempo, e anche nel corso di una sola conversazione.|...| Gli obiettivi sono diversi, e restano non specificati. Non c'è da stupirsi che ci voglia tanto per giungere a compromessi.

Chiaramente, gli obiettivi devono essere fissati, e quindi i modelli mentali anche, altrimenti come facciamo a determinare gli operatori (per usare il linguaggio *GPS*) da applicare agli oggetti che vogliamo trasformare in «oggetti desiderati»? E la confusione dei modelli mentali, osserva Forrester, è in larga misura dovuta alla confusione dello stesso linguaggio umano. Anche a questo bisogna porre rimedio.

> I modelli per il computer differiscono dai modelli mentali in modi importanti. I modelli per il computer sono specificati in modo esplicito. Le notazioni «matematiche» usate per descrivere il modello sono prive di ambiguità. È un linguaggio più chiaro e preciso dei linguaggi parlati, come inglese o francese. Il linguaggio dei modelli per il computer è più semplice. Il suo vantaggio sta nella chiarezza del significato e nella semplicità della sintassi. Il linguaggio di un modello per il computer può essere compreso quasi da chiunque, indipendentemente dal grado di istruzione. Per di più, qualsiasi concetto e connessione che possano essere espressi in modo chiaro nel linguaggio ordinario possono venire tradotti nel linguaggio della modellistica per il computer.

C'è da restare stupiti del fatto che il linguaggio ordinario, con tutte le sue caratteristiche anti-funzionali, sia sopravvissuto finora. E se è così chiaro che qualsiasi concetto e connessione possono essere tradotti in termini da computer, perché linguisti come Halle, Jakobson, Chomsky continuano a lottare così vigorosamente? E perché esistono ancora dei poeti? Per restare al nostro argomento attuale, comunque, semplicemente non è vero che «quasi chiunque» possa capire il linguaggio, mettiamo, dei modelli per computer di Forrester. Questi ultimi sono stati diffusamente accettati principalmente perché prodotti da uno scienziato famoso affiliato a un'università prestigiosa, e perché i loro risultati sono «quello che dice il computer». La maggior parte dei ministri, dei leader sindacali e dei commentatori sociali che hanno partecipato al dibattito sui «limiti dello sviluppo» non era in grado di leggere i programmi che stavano alla base della controversia, più di quanto sia in grado di leggere le equazioni della fisica quantistica. Tuttavia, come l'ammiraglio Moorer, essi trovavano utile «fidarsi» della macchina.

Il professor Forrester, infine, rassicura il suo pubblico sul fatto che «sono in vista i mezzi» per cominciare a porre fine all'incertezza (riesce a vederli lui, naturalmente):

> La grande incertezza dei modelli è la loro incapacità di prevedere le conseguenze delle interazioni tra le parti di un sistema. Questa incertezza è completamente eliminata nei modelli per il computer. Dato un insieme di ipotesi, il computer delinea le conseguenze che ne risultano senza dubbi o errori.

Egli prosegue dicendo che, sebbene nel nostro sistema sociale non vi sia «alcuna utopia», né alcun modo sostenibile di comportarsi, che sia libero da pressioni e da *stress*, ci sono alcuni possibili modi di comportamento che sono più «desiderabili» di altri. E come sono resi possibili questi modi di comportamento più desiderabili?

> Sembra che essi siano possibili soltanto se abbiamo una buona comprensione della dinamica dei sistemi, e la volontà di sopportare l'auto-disciplina e le pressioni che devono accompagnare i modi desiderabili.

C'è senz'altro qualche accezione delle parole «sistema» e «dinamica» che darebbe un senso benigno a questa considerazione. Ma nel contesto in cui queste parole sono state pronunciate, esse hanno il particolare significato conferito loro da Forrester. È allora chiaro che il messaggio di Forrester è più o meno lo stesso di Skinner e degli altri: il solo modo per acquisire la comprensione che sola può portare ai «modi di comportamento desiderabili» è usare i metodi di «analisi scientifica» alla Forrester (o alla Skinner, o alla *GPS* ecc.).

I diversi sistemi e programmi che abbiamo discusso hanno delle caratteristiche molto significative in comune: sono tutti, in un certo senso, semplici; tutti distorcono il linguaggio e ne abusano; e tutti, mentre smentiscono di avere un contenuto normativo, appoggiano un autoritarismo basato sulla competenza. Questo appoggio è, naturalmente, camuffato dal loro uso della retorica, adagiata in un linguaggio apparentemente neutro, carico di espressioni gergali, fattuale (cioè, da quelle che l'uomo della strada chiama «stronzate»). Queste caratteristiche in comune sono, in una certa misura, separabili, ma non sono indipendenti una dall'altra.

Gli aspetti più superficiali della semplicità di questi sistemi — riflessi nell'immagine semplicistica che danno dei loro argomenti — sono immediatamente visibili. Simon, per esempio, vede l'uomo come un organismo «piuttosto semplice». La complessità «apparente» del suo comportamento è dovuta alla complessità del suo ambiente. In ogni caso, l'uomo può venire simulato da un sistema sensibile soltanto a «pochi parametri semplici», un sistema che consiste soltanto di pochi (certamente molto meno di, mettiamo, diecimila) «processi informatici elementari». Il direttore di laboratorio che ho citato crede che i problemi dell'intenzionalità umana possano essere affrontati in modo utile facendo ricerche sulle tecniche di debugging dei programmi per computer, e questa fede è condivisa da molti suoi colleghi. Skinner vede l'uomo essenzialmente come un prodotto passivo (una vittima) del suo patrimonio genetico, e della

sua storia fatta di contingenze che rafforzano quel patrimonio. La differenza principale tra il sistema di Skinner e il *GPS* sembra essere il fatto che Skinner vuole guardare soltanto al «comportamento di *input/output*», mentre gli artefici del *GPS* e dei sistemi analoghi sentono di poter dire qualcosa anche su ciò che avviene dentro l'organismo. Ma le differenze filosofiche tra i due atteggiamenti sono modeste. Forrester vede letteralmente tutto il mondo in termini di cicli di retroazione.

> I cicli di retroazione sono gli elementi costitutivi fondamentali dei sistemi. |...| Un ciclo di retroazione è composto da due tipi di variabili, che chiamiamo qui variabili di proporzione e di livello. Questi due tipi di variabili sono necessari e sufficienti |...| Le variabili di livello sono accumulazioni o integrazioni |...| Le proporzioni del flusso fanno cambiare i livelli. I livelli forniscono i dati d'ingresso per le equazioni di proporzione che controllano i flussi. Le equazioni di proporzione sono le dichiarazioni della politica del sistema. Esse determinano il modo in cui l'informazione disponibile è convertita in corrente di azione. |...| Un'equazione di proporzione stabilisce la discrepanza tra l'obiettivo e la situazione osservata. E infine, l'equazione di proporzione determina l'azione che risulterà da quella discrepanza [15].

Notate la sovrapposizione di linguaggio tra questa affermazione e la discussione condotta da Newell e Simon sui problemi e la loro risoluzione (capitolo 6). Nella seconda si parla di «oggetto attuale» e «oggetto desiderato», delle differenze tra i due, degli operatori che riducono queste differenze, e così via. La differenza tra il loro sistema e quello di Forrester sta principalmente nel diverso insieme di concetti primitivi della «elaborazione elementare di informazioni» che ciascuno usa e, naturalmente, nel fatto che il *GPS* usa metodi euristici per ridurre la ricerca degli operatori ecc., mentre nel sistema di Forrester tutto è esplicitamente algoritmizzato in termini di variabili di proporzione e di livello, avviluppate in cicli di retroazione. Ma la visione del mondo rappresentata da questi due sistemi è fondamentalmente la stessa. Ed è molto semplice.

Ma questi sistemi sono semplici anche in un senso più profondo e più importante. Essi hanno ridotto la ragione stessa al solo ruolo che essa svolge nel dominio sulle cose, sull'uomo e infine sulla natura.

> I concetti sono stati ridotti a riassunti delle caratteristiche che parecchi campioni hanno in comune. Denotando somiglianza, i concetti eliminano la seccatura di dover enumerare delle qualità, e in questo modo funzionano meglio nell'organizzazione della materia di conoscenza. Sono considerati semplici abbreviazioni degli articoli cui si riferiscono. Qualsiasi uso che vada al di là di un riassunto tecnico, ausiliario, di dati di fatto è stato eliminato come fosse un'ultima traccia

[15] J.W. Forrester, *Urban Dynamics* («Dinamiche urbane»), The MIT Press, Cambridge, Mass. 1969, pp. 1311.

di superstizione. I concetti sono diventati apparecchi «semplificati», razionalizzati per risparmiare fatica |...| Il pensiero stesso è stato ridotto al livello di procedimento industriale |...| in breve, è stato reso parte e lotto di produzione [16].

Chi non conosce la base tecnica dei sistemi che abbiamo discusso fin qui non potrà apprezzare l'agghiacciante accuratezza del loro resoconto dato in questo passaggio. Esso è stato scritto dal filosofo-sociologo Max Horkeimer nel 1947, molti anni prima che le forze che già allora stavano eclissando la ragione (per usare un'espressione di Horkeimer stesso) venissero letteralmente incarnate da macchine.

Questo passaggio, specialmente in considerazione di quando e da chi fu scritto, ci mostra ancora una volta che il computer, come viene usato attualmente dall'*élite* tecnologica, non è una causa di nulla. È piuttosto uno strumento spremuto al servizio della razionalizzazione, del sostegno e del supporto alla componente ideologica più conservatrice, anzi reazionaria, dello *Zeitgeist* attuale.

Come si vede così chiaramente nei diversi sistemi in esame, il significato è stato completamente trasformato in funzione. Il linguaggio, e quindi anche la ragione, è stato trasformato in nient'altro che uno strumento per influenzare le cose e gli avvenimenti del mondo. Niente di ciò che questi sistemi fanno ha un significato intrinseco. Ci sono soltanto obiettivi dettati da onde di marea cui non ci si può opporre. Ci sono soltanto analisi mezzi-fini per rivelare le discrepanze tra il modo in cui le cose sono, la «situazione osservata», e il modo in cui il destino che ci è toccato ci dice che desideriamo siano. Nel processo di adattamento a questi sistemi, noi, e perfino quelli di noi che sono ammiragli, abbiamo non solo castrato noi stessi (vale a dire, ci siamo rassegnati all'impotenza), ma anche il nostro linguaggio. Perché ora quel linguaggio è diventato soltanto un altro strumento; tutti i concetti, le idee, le immagini che artisti e scrittori non sanno parafrasare in linguaggi comprensibili al computer hanno perso la loro funzione e la loro potenza. Questo ce lo dice nel modo più chiaro Forrester (ma è evidente che anche gli altri annuiscono): «Qualsiasi concetto e connessione che possano essere espressi in modo chiaro nel linguaggio ordinario possono venire tradotti nel linguaggio della modellistica per il computer». L'onere di provare che qualcosa è stato «espresso in modo chiaro» è sul poeta. Non c'è da stupirsi, data questa visione del linguaggio, che la distinzione tra vivente e senza vita, tra uomo e macchina, sia diventata qualcosa di non proprio reale, al massimo una questione di sfumature!

La corruzione del linguaggio è parte molto integrante della retorica dell'*élite* tecnologica. Abbiamo già fatto notare la trasformazione attuata da Minsky della parola «capire» in un termine puramente strumentale. Ed è questa sua interpretazione che, ovviamente, permea tutti i sistemi che abbiamo discusso.

[16] M. Horkeimer, *Eclipse of Reason* («L'eclisse della ragione»), Seabury, New York 1974, p. 21. Questo importante libro è stato pubblicato per la prima volta dalla Oxford University Press, N.Y., nel 1947.

L'uso fatto da Newell e Simon della parola «problema» è un altro esempio, e altrettanto significativo.

Nel periodo critico passato anni fa dai *campus* universitari americani, si sentivano spesso oratori bene intenzionati sostenere che la tensione, almeno nei loro *campus*, era causata principalmente da una comunicazione inadeguata tra le diverse componenti dell'università, per esempio tra corpo insegnante, amministrazione, studenti, personale non docente. Il «problema» era perciò visto fondamentalmente come un problema di comunicazione, e quindi tecnico. Era perciò risolubile con mezzi tecnici, come per esempio la creazione di varie «linee calde» con, mettiamo, l'ufficio del presidente o del rettore. Forse c'erano davvero delle difficoltà di comunicazione; è normale che ce ne siano in quasi tutte le università. Ma questa visione del «problema» — del tutto coerente con la visione della «risoluzione umana dei problemi» di Newell e Simon, e con il ragionamento strumentale — nasconde, seppellisce attivamente l'esistenza dei veri conflitti. Può darsi, per esempio, che gli studenti abbiano degli interessi etici, morali, politici genuini in conflitto con quelli che l'amministrazione sente di avere, e che ciascuna componente capisca benissimo gli interessi dell'altra. Allora esiste un problema vero e proprio, non soltanto una difficoltà di comunicazione, e certamente non un problema che si possa risolvere con l'espediente tecnico delle linee calde. Ma la ragione strumentale trasforma qualsiasi dilemma, per quanto genuino, in un semplice paradosso che può quindi essere sciolto applicando la logica, il calcolo. Tutti gli interessi in conflitto sono rimpiazzati dagli interessi della sola tecnica.

Anche questa, come la storia di Philip Morrison, è una parabola. Il suo senso più ampio è che la contaminazione della parola «problema» ha portato con sé la mistica del *problem solving*, la risoluzione dei problemi, con effetti catastrofici sul mondo interno. Quando ogni problema sulla scena internazionale è visto dai «migliori e più brillanti» risolutori di problemi come un problema puramente tecnico, guerre come la guerra del Vietnam diventano davvero inevitabili. Il riconoscimento di interessi autenticamente contrastanti, ma legittimi, in società coesistenti — riconoscimento che è certamente una precondizione per la risoluzione e la sistemazione dei conflitti — viene reso impossibile fin dall'inizio. Si usano invece i criteri più semplici per rivelare le differenze, per cercare il modo di ridurle, e infine per applicare operatori che trasformino gli «oggetti attuali» negli «oggetti desiderati». È infatti del tutto ragionevole, se «ragione» significa ragione strumentale, applicare la forza militare americana, i B-52, il napalm e tutto il resto, al Vietnam «dominato dai comunisti» (chiaramente un «oggetto indesiderabile»), come «operatore» per trasformarlo in un «oggetto desiderabile», vale a dire in una nazione al servizio degli interessi americani.

La meccanizzazione della ragione e del linguaggio ha conseguenze che vanno ben al di là di qualsiasi effetto previsto dai risolutori di problemi che abbiamo citato. Horkeimer, molto tempo prima che il computer diventasse un feticcio e desse forma concreta all'eclisse della ragione, ci diede la prospettiva necessaria:

> Giustizia, uguaglianza, felicità, tolleranza, tutti i concetti che |...| nei secoli passati erano ritenuti intrinseci alla, o sanciti dalla, ragione, hanno perduto le loro radici intellettuali. Sono ancora obiettivi e fini, ma non c'è più alcuna entità razionale autorizzata a valutarli e a collegarli a una realtà obiettiva. Appoggiati come sono da documenti storici venerabili, possono ancora godere di un certo prestigio, e alcuni sono contenuti nella legge suprema delle nazioni più grandi. Ciò nonostante, manca loro ogni conferma da parte della ragione nella sua accezione moderna. Chi può dire che uno qualunque di questi ideali sia legato alla verità più strettamente del suo contrario? Secondo la filosofia dell'intellettuale medio dei nostri tempi, c'è una sola autorità, e cioè la scienza, concepita come classificazione dei fatti e calcolo delle probabilità. L'affermazione che giustizia e libertà siano di per sé meglio che ingiustizia e oppressione è, dal punto di vista scientifico, non verificabile e inutile. Ha finito per suonare tanto priva di senso quanto lo sarebbe l'affermazione che il rosso è più bello del blu, o che un uovo è meglio del latte [17].

Come abbiamo osservato anche noi, il realizzarsi di sistemi complessi che non hanno autore, di cui sappiamo solo che ci sono stati dati in qualche modo dalla scienza, e che parlano con la sua autorità, non consente di fare alcuna domanda su verità o giustizia.

Io non so dire perché le persone che ho citato vogliano far diventare realtà gli sviluppi che predicono. Alcuni di loro mi hanno detto che ci lavorano sopra per la ragione — fallimentare dal punto di vista morale — che «se non lo facciamo noi, lo farà qualcun altro». Hanno paura che gente malvagia sviluppi macchine super-intelligenti e le usi per dominare l'umanità, e che l'unica difesa contro queste macchine nemiche saranno delle macchine super-intelligenti controllate da noi, cioè da persone ben intenzionate. Altri confessano di avere abdicato alla loro autonomia, rifacendosi al «principio» dell'ineluttabilità tecnologica. Ma, alla fine dei conti, tutto quello che posso dire con sicurezza è che queste persone *non* sono stupide. Tutto il resto è un mistero.

È appena più facile capire come mai l'opinione pubblica accolga le loro idee almeno con serenità, e talvolta perfino con entusiasmo. La retorica dell'intellighenzia tecnologica può essere attraente, perché appare come un invito a ragionare. E lo è, infatti. Ma, come ho dimostrato, stimola ragionamenti strumentali, e non la razionalità umana autentica. Pubblicizza risposte facili, e giustificate «scientificamente», a qualunque problema si possa concepire. Sfrutta il mito della competenza. Anche qui, la contaminazione del linguaggio svolge un ruolo importante. Il linguaggio dell'intellighenzia artificiale, dei modificatori del comportamento, e degli ingegneri dei sistemi è mistificatorio. Le persone, le cose, gli avvenimenti sono «programmati», si parla di «input» e «output», di cicli di retroazione, di variabili, parametri, processi e così via, finché ci si astrae completamente da qualsiasi contatto con le situazioni concrete. Allora restano soltanto grafici, insiemi di dati, stampati del calcolatore. E soltanto «noi», gli esperti, li possiamo capire. «Noi» dimostriamo — anche

[17] *Ivi*, pp. 23-24.

se soltanto per avere buone relazioni pubbliche — la nostra preoccupazione per le conseguenze sociali delle «nostre» azioni e dei «nostri» piani. Documenti programmatici come quello che ho citato hanno quasi sempre un paragrafo iniziale che, di passaggio, fa riferimento al potenziale di distruzione dei nostri strumenti. E «noi» scriviamo saggi sulle implicazioni sociali dei nostri congegni. Ma, come ho osservato altrove, alla fine questi lavori risultano essere fatti a uso e consumo soltanto nostro.

> La struttura del tipico saggio su «L'impatto del computer sulla società» è come segue: dapprima c'è un'affermazione che comincia con «Da una parte...». Racconta tutte le belle cose che i computer hanno già fatto per la società, e spesso tenta perfino di dimostrare che l'ordine sociale sarebbe già crollato se non fosse per la «rivoluzione del computer». Di solito segue una nota di cautela tipo «D'altra parte...», che racconta alcuni problemi che l'introduzione del computer porta con sé. Di solito vengono citati la minaccia portata alla *privacy* individuale dalle grandi banche-dati, e il pericolo di disoccupazione su larga scala indotto dall'automazione dei processi industriali. Infine, si applaude ai gloriosi risultati presenti e futuri del computer, mentre si mostra che i pericoli cui si era accennato nella seconda parte possono venire alleviati usando sofisticati rimedi tecnologici. Il paragrafo conclusivo consiste in una supplica per un generoso sostegno della società a ulteriori e più estese ricerche e sviluppi sui computer. Questa di solito viene associata all'affermazione più o meno sottile che soltanto la scienza dei computer, e quindi soltanto lo scienziato dei computer, può proteggere il mondo dalle ricadute, a parere di tutti pericolose, della tecnologia dei computer applicata [18].

Il vero messaggio degli scritti di questo tipo è perciò che l'esperto si prenderà cura di tutto, anche dei problemi creati da lui stesso. Ha bisogno di più denaro. Quello, sempre. Ma egli si affanna a rassicurare un pubblico che, comunque, non vuole sapere.

E qual è la risposta del tecnologo ad accuse come quelle fatte qui? Per prima cosa, esse vengono lasciate perdere in quanto puramente filosofiche. Per esempio, il mio articolo sull'impatto del computer sulla società attirò centinaia di lettere, ma soltanto una da un membro della comunità dell'intelligenza artificiale. Venne da un ex-studente del professor Simon, e diceva tra l'altro:

> Per quanto riguarda la società nel suo complesso, gli effetti primari della tecnologia dei computer sono più importanti dei loro (sic) effetti collaterali. Soltanto la persona più incline alla filosofia troverà più importanti i potenziali effetti collaterali. [...] Ci vuole una persona rara per passare più di qualche ora a ponderare le implicazioni filosofiche [19].

[18] J. Weizenbaum, *On The Impact of Computers on Society* («Sull'impatto dei computer sulla società»), «Science», vol. 176, 1972, nr. 12, pp. 609-614.

[19] L.S. Coles, *Lettera al Direttore,* «Science», 10 novembre 1972.

Questo, naturalmente, è del tutto in accordo con l'osservazione di Horkeimer secondo cui il linguaggio ha perso perfino il diritto di parlare in termini non strumentali, cioè filosofici. Ma una risposta più direttamente significativa venne dal dottor Kenneth B. Clark durante un simposio tenutosi al MIT non molto tempo fa. Egli aveva appena espresso il suo disappunto per il fatto che il MIT non dedicasse una parte maggiore delle sue risorse alla soluzione dei problemi sociali. Disse (cito a memoria): «Questo è un grande istituto dedicato a scienza e tecnologia, ed esperto in esse. Perché, in questo tempo di angoscia, non applicate i vostri strumenti, le vostre tecniche agli scottanti problemi sociali di oggi?».

Io suggerii che forse le risposte ai problemi urgenti della nostra epoca non si potevano trovare esclusivamente nella scienza e nella tecnologia. Dissi che la sua stessa ricerca di soluzioni tecnologiche ai grandi problemi, per esempio la sua proposta di somministrare regolarmente dei tranquillanti ai leader mondiali [20], poteva essere sostanzialmente fuorviante. Egli rispose: «Sono giunto molto tempo fa alla conclusione che le risposte ai grandi problemi che l'uomo ha di fronte in ogni epoca possono venire soltanto dal pensiero razionale. La sola alternativa è quel tipo di irragionevolezza che, come abbiamo visto, porta soltanto a violenza e distruzione».

Si potrebbe benissimo sottoscrivere questo punto di vista, ma soltanto se per razionalità si intende qualcos'altro che non la pura applicazione di scienza e tecnologia, se la razionalità non è considerata automaticamente, e per implicazione, equivalente a calcolabilità e logicità. L'alternativa a quel tipo di razionalità che vede nella psicotecnologia la soluzione ai problemi del mondo non è l'irragionevolezza. È invece la ragione restaurata a dignità umana, ad autenticità, a stima di sé e ad autonomia individuale.

La ragione strumentale ha fatto delle parole un feticcio circondato da un'atmosfera da magia nera. E soltanto i maghi hanno i diritti che spettano agli iniziati. Soltanto loro possono dire cosa significhino le parole. Ed essi giocano con le parole e ci ingannano. Quando Skinner oppone la scienza al buon senso comune, e sostiene che la scienza è molto superiore, egli intende la sua «scienza comportamentale», e intende «comune» — in «buon senso comune» — come un peggiorativo. Non intende quel buon senso comune ispirato da una visione culturale condivisa da molti, né un buon senso comune che, senza proprio nessuna «ragione», recalcitra di fronte all'idea che libertà e dignità siano concetti assurdi e fuori moda.

Il tecnologo ribatte sempre che opinioni come quelle espresse qui sono anti-tecnologiche, anti-scientifiche e in ultima istanza anti-intellettuali. Egli cercherà di interpretare tutte le tesi contrarie alle sue visioni megalomaniache come tesi miranti all'abbandono della ragione, della razionalità, della scienza e della tecnologia, e favorevoli alla pura intuizione, alle sensazioni, all'irragionevolezza indotta dalla droga, e così via. In realtà, io sto sostenendo proprio

[20] K.B. Clark, *Leadership and Psychotechnology* («Leadership e psicotecnologia»), «New York Times», 9 novembre 1971.

la razionalità. Ma sostengo che essa non può venire separata da intuizioni e sensazioni. Io sostengo l'uso razionale di scienza e tecnologia, non la loro mistificazione, né tanto meno il loro abbandono. Io incito all'introduzione del pensiero etico nella pianificazione scientifica. Combatto l'imperialismo della ragione strumentale, non la ragione.

Si dice che vi siano sempre stati uomini che pensavano che il loro tempo fosse pieno dei più grandi presagi di catastrofi a venire, e addirittura che la loro epoca fosse la peggiore di tutta la storia dell'umanità. Certamente, noi che eravamo vivi e svegli quando sembrava che il fascismo fosse ovunque vittorioso, vedevamo prepararsi la fine della civiltà stessa. In qualche modo, la civiltà è sopravvissuta a quella minaccia, una minaccia che i giovani di oggi non possono più capire. Ma non si può dire che la civiltà sia sopravvissuta completamente intatta a questa minaccia, o alla Grande Guerra che la precedette di due soli decenni. Venimmo a sapere, come mai prima, cosa l'uomo possa fare al suo prossimo. La Germania intraprese la «soluzione finale» al suo «Problema ebraico» come un esercizio da libro di testo sul ragionamento strumentale. L'umanità rabbrividì per un attimo, quando non poté più distogliere lo sguardo da quanto era successo, quando cominciarono a circolare le fotografie scattate dagli stessi assassini, e quando i miserevoli superstiti riemersero alla luce. Ma alla fine, non fece alcuna differenza. La stessa logica, la stessa applicazione fredda e crudele della ragione calcolatrice, portarono nei venti anni successivi al macello di almeno altrettante persone quante ne erano state vittime dei tecnici del Reich millenario. Non abbiamo imparato nulla. La civiltà è in pericolo oggi come lo era allora.

Ma se ogni epoca ha ascoltato lo stesso lamento di Cassandra, allora ogni epoca ha appreso anche quanto poco profetico esso sembri sempre rivelarsi. Sono state distrutte delle civiltà, e molte. Ma mai l'umanità. Ma questa volta è diverso. Siamo stufi di sentirlo dire, ma non possiamo negarlo: questa volta l'uomo è davvero in grado di distruggere tutto. Soltanto una sua decisione può salvarlo.

Si usava anche dire che la religione fosse l'oppio dei popoli. Immagino che quel modo di dire significasse che i popoli erano drogati dalle visioni della bella vita che sarebbe sicuramente spettata loro, se soltanto avessero sopportato pazientemente l'inferno terreno che i loro padroni avevano creato. D'altra parte, può darsi che la religione non dia affatto assuefazione. Se l'avesse data, forse Dio non sarebbe morto, e la nuova razionalità non avrebbe sbaragliato la pietà. Ma la ragione strumentale, la tecnica trionfante, la scienza sfrenata, queste sì che danno assuefazione. Creano una realtà concreta, un incubo che si realizza da sé. Può ancora essere che il tecnologo ottimista abbia ragione: forse abbiamo raggiunto il punto di non ritorno. Ma perché l'equipaggio che ci ha portato fin qui esulta? Perché i passeggeri non alzano lo sguardo dai loro giochi? E infine, ora che siamo noi, e non più Dio, a giocare a dadi con l'universo, come possiamo evitare di farci sbancare?

10
CONTRO L'IMPERIALISMO DELLA RAGIONE STRUMENTALE

L'opinione secondo cui l'uomo si è procacciato con scienza e tecnologia una potenza enorme è una frase fatta così grossolanamente banale che, paradossalmente, benché tutti continuino a crederla vera, viene ripetuta sempre meno di frequente nelle conversazioni serie. Il paradosso sorge perché una frase fatta che smette di essere luogo comune, smette anche di essere considerata una frase fatta. In alcuni circoli, anzi, dopo che non la si è sentita per un po', essa può anche venir percepita in modo opposto, vale a dire come una profonda verità. Anche in questo si cela una parabola: la potenza che l'uomo ha acquistato con la sua scienza e la sua tecnologia è stata trasformata in impotenza.

La gente comune certamente se ne rende conto. Studs Terkel, in uno studio monumentale sul lavoro quotidiano negli Stati Uniti, scrive che:

> La massa non nasconde il suo scontento. |...| «Sono una macchina», dice il saldatore. «Sono ingabbiato», dice l'impiegato di banca allo sportello, e il portiere d'albergo echeggia. «Sono un mulo», dice l'operaio all'altoforno. «Una scimmia potrebbe fare quello che faccio io», dice l'impiegato alla *reception*. «Sono meno di un attrezzo da officina», dice il lavoratore emigrante. «Sono un oggetto», dice la modella d'alta moda. I colletti blu e quelli bianchi ricorrono alla stessa espressione: «Sono un robot» [1].

Forse la gente comune pensa che, sebbene essa sia impotente, un potere esista, e cioè quello esercitato dai suoi capi. Ma abbiamo visto che il Segretario di Stato americano ritiene che gli avvenimenti ci «accadano» semplicemente, e che il capo dello *staff* presidenziale americano confessa di essere diventato schiavo del computer. Neanche i nostri capi sanno trovare il potere.

Perfino i medici, una volta simbolo di potere di una cultura, sono impotenti nel loro diventare sempre più un anello di congiunzione tra i loro pazienti e le principali case farmaceutiche. I pazienti, a loro volta, sono sempre più dei semplici oggetti passivi, cui si applicano cure e si fanno interventi. Le loro ri-

[1] S. Terkel, *Working* («Lavorare»), Pantheon, New York 1974, p. XI.

sorse curative interiori, le loro capacità di auto-reintegrazione, sia psichiche che fisiche, sono considerate sempre più come irrilevanti da una medicina che riconosce a stento la differenza tra un paziente umano e un oggetto prodotto alla catena di montaggio. Il movimento, oggi in ascesa, della bio-retroazione *(bio-feedback)* può forse essere il penultimo atto della tragedia che separa l'uomo dalla natura; l'uomo non percepisce neppure più se stesso, il suo corpo, in modo diretto, bensì soltanto tramite lancette, spie luminose e suoni ronzanti prodotti da strumenti che gli vengono attaccati come tachimetri a un'automobile. L'ultimo atto della tragedia è, naturalmente, l'olocausto finale che spazza via del tutto la vita.

L'ineluttabilità tecnologica può così essere vista come un solo elemento di una sindrome molto più estesa. La scienza ha promesso all'uomo la potenza. Ma, come succede così spesso quando le persone vengono sedotte da promesse di potenza, il prezzo preteso in anticipo e durante tutto il cammino, e il prezzo effettivamente pagato, è la schiavitù e l'impotenza. La ragione strumentale può prendere decisioni, ma c'è una bella differenza tra decidere e scegliere.

Le persone di cui parla Studs Terkel prendono decisioni tutto il giorno, ogni giorno. Ma non sembra che facciano delle scelte. Sono, come testimoniano esse stesse, come il robot di Winograd. Gli si chiede «Perché l'hai fatto?», e risponde «Perché questa o quella diramazione decisionale nel mio programma ha dato quel risultato». E allora si chiede «Perché sei arrivato a quella diramazione?», e risponde di nuovo allo stesso modo. Ma la sua risposta finale è: «Perché tu mi hai detto di farlo». Forse, ogni azione umana implica una catena di calcoli da effettuare a quelli che un analista di sistemi chiamerebbe «nodi decisionali». Ma la differenza tra un'azione meccanica e una autenticamente umana è che la seconda termina in un nodo il cui parametro decisivo non è «Perché tu mi hai detto di farlo», bensì «Perché ho scelto di farlo». A quel punto, i calcoli e le spiegazioni vengono rimpiazzati dalla verità. Anche qui, si rivela la povertà dell'ipotesi di Simon secondo cui:

> L'uomo nella sua interezza, come la formica, se visto come un sistema nell'atto di comportarsi, è piuttosto semplice. L'apparente complessità del suo comportamento in un periodo di tempo sufficientemente lungo è in gran parte un riflesso della complessità dell'ambiente in cui si trova.

Perché quest'ipotesi fosse vera, dovrebbe anche essere vero che la capacità di scegliere dell'uomo è limitata come quella della formica, che l'uomo non ha più volontà e scopo, né — cosa forse più importante — un maggiore senso di impegno nei confronti di se stesso come parte del continuo naturale, rispetto alla formica. Ancora una volta, è un mistero che qualcuno possa credere che sia questa la vera situazione dell'uomo.

Ma, di tanto in tanto, sembra che una piccola luce penetri nella fitta nebbia che oscura le autentiche capacità dell'uomo. Di recente, per esempio, un gruppo di illustri biologi ha invitato dei colleghi a cessare certi esperimenti,

in cui vengono creati dei nuovi tipi di plasmidi batterici biologicamente funzionali[2]. Essi hanno espresso «seria preoccupazione che alcune di queste molecole artificiali di DNA ricombinate potrebbero rivelarsi biologicamente pericolose». La loro preoccupazione, scrivono, riguarda «le possibili conseguenze disgraziate dell'applicazione indiscriminata di queste tecniche». Il loro è certamente un passo nella direzione giusta, e la loro iniziativa va applaudita. Eppure, ci si può domandare, perché ritengono di dover dare una giustificazione per quello che raccomandano di fare? Forse che il dovere primario degli uomini, compresi gli uomini di scienza, di esentare la vita stessa dalla follia di trattare tutto come un oggetto, non è una ragione sufficiente, anzi una ragione che non ha neppure bisogno di essere pronunciata esplicitamente? Perché bisogna spiegarla? Si direbbe che perfino i gesti più nobili delle persone meglio intenzionate siano avvelenati dal clima corrosivo dei valori del nostro tempo.

Una facile spiegazione di questo, e forse contiene della verità, è che le buone intenzioni abbiano soppiantato del tutto la nobiltà d'animo. Ma ce n'è una più sottile. Il nostro tempo si vanta di aver raggiunto la libertà dalla censura per cui hanno lottato i libertari di tutte le epoche. Oggi è possibile discutere di questioni sessuali più liberamente che mai prima, le donne stanno incominciando a trovare la posizione che loro spetta nella società e, in generale, idee che potevano solo essere sussurrate fino a un decennio fa o poco più, possono ora circolare senza restrizioni. Il merito di queste grandi conquiste viene reclamato dal nuovo spirito di razionalismo, un razionalismo che — si sostiene — è riuscito infine a strappare dagli occhi dell'uomo i veli imposti dal pensiero mistico, dalla religione, e da illusioni potenti come libertà e dignità. La scienza ci ha dato questa grande vittoria sull'ignoranza. Ma a un esame più accurato anche questa vittoria può essere vista come il trionfo orwelliano di un'ignoranza ancora maggiore: quello che abbiamo ottenuto è un nuovo conformismo, che ci permette di dire qualsiasi cosa che possa essere detta nei linguaggi funzionali della ragione strumentale, ma ci proibisce di alludere a quella che Ionesco chiamava «verità vivente». Proprio come i nostri schermi televisivi ci possono mostrare in «colore vivente» una violenza sfrenata, ma non scene di autentico amore intimo — la prima, con un'inversione di valori di per sé oscena, viene considerata «realistica», mentre le seconde sono definite oscene —, così ci è concesso di discutere la creazione della vita e la sua manipolazione «oggettiva», ma non possiamo citare Dio, la grazia o la moralità. Forse, i biologi che invitano i loro colleghi a fare la cosa giusta, ma per le ragioni sbagliate, sono in realtà motivati dal profondo rispetto per la vita e dalla loro autentica umanità, soltanto non osano dirlo. In ogni caso, argomenti di questo tipo non sarebbero «efficaci», vale a dire, strumentali.

Se è così, allora coloro che censurano le loro stesse parole lo fanno — per usare un'espressione non più di moda — a rischio della loro anima.

[2] P. Berg et al., lettera a «Science», 26 luglio 1974.

C'è anche un altro modo per giustificare la rinuncia di uno scienziato a un particolare indirizzo di ricerca; ed è un modo da cui tutti noi possiamo ricavare lezioni pertinenti per la nostra vita personale. Parte del principio secondo cui il campo delle responsabilità che uno ha deve essere commensurabile con il campo degli effetti delle sue azioni. Nel passato, questo principio ha portato a un sistema etico che riguardava principalmente il modo in cui le persone si comportavano una nei confronti dell'altra. I comandamenti biblici, per esempio, parlano principalmente di quelli che sono i doveri di un individuo nei confronti della sua famiglia e del suo vicinato. In epoca biblica, pochi erano in grado di fare qualcosa che potesse influenzare persone al di là del loro spazio vitale. La scienza e la tecnologia dell'uomo hanno modificato in modo radicale questa situazione. Non soltanto le azioni dell'uomo moderno possono influenzare l'intero pianeta che è il suo habitat, ma possono determinare il futuro dell'intera specie umana. Ne segue perciò che l'uomo, e in particolare l'uomo scienziato e ingegnere, ha responsabilità che vanno al di là della sua situazione immediata, e che di fatto si estendono direttamente alle generazioni future. Queste responsabilità sono particolarmente gravose, per il fatto che le generazioni future non possono sostenere la loro causa oggi. Tutti noi siamo loro amministratori fiduciari [3].

La pubblica rinuncia dei biologi, comunque essi la giustifichino, è un esempio che tutti gli scienziati devono imitare. Questo significa forse suggerire che gli scienziati dovrebbero chiudere la loro mente a una certa classe di ipotesi «immorali»? Niente affatto. Un'ipotesi scientifica, almeno dal punto di vista scientifico, o è vera oppure è falsa. Questo vale, per esempio, per le ipotesi di Simon secondo cui l'uomo è «piuttosto semplice», e può venire interamente simulato da una macchina, così come per l'ipotesi di McCarthy secondo cui esiste un sistema di calcolo logico in termini del quale è possibile formalizzare tutta la realtà. Sarebbe uno sciocco errore di logica etichettare queste ipotesi (o qualunque altra) come morali o immorali oppure, se si vuole, come responsabili o irresponsabili.

Ma benché un'ipotesi scientifica non possa avere di per sé alcuna dimensione morale o etica, la decisione di un individuo di adottarla, anche in via sperimentale (per non parlare della decisione di annunciare la propria fede in essa all'opinione pubblica), sicuramente coinvolge giudizi di valore, e ha perciò quelle dimensioni. Come ha scritto di recente l'economista di Harvard Marc J. Roberts:

> Immaginiamo di dover scegliere tra due ipotesi. Indipendentemente da quella che scegliamo, c'è sempre la possibilità che sia corretta l'altra. Ovviamente, la probabilità relativa di fare un errore, quando scegliamo l'una o l'altra, ha impor-

[3] La tesi di questo capoverso è sostenuta in modo molto più diffuso, e più persuasivo, da Hans Jonas nel suo articolo *Technology and Responsibility: Reflections on the New Tasks of Ethics* (Tecnologia e responsabilità: riflessioni sui nuovi compiti dell'etica), apparso in «Social Research», vol. 40, 1973, nr. 1, pp. 31-34. Sono grato al professor Langdon Winner del MIT per aver richiamato la mia attenzione su quel lavoro.

tanza; ma lo stesso vale per il costo che avrebbe ciascuno dei due errori, il costo di supporre A quando in realtà è vero B, e viceversa. Potremmo benissimo scegliere di rischiare un piccolo costo più probabile, anziché uno grande meno probabile. Eppure, le dimensioni del costo di uno sbaglio nei due casi non possono essere determinate se non sulla base dei nostri valori.

Considerate un esempio estremo: l'opinione che vi siano differenze genetiche nel funzionamento mentale delle diverse razze umane. Supponiamo che la società dovesse accettare quest'opinione, ed essa si dimostrasse falsa. Io credo che si sarebbe fatto del male enorme. D'altra parte, supponiamo che la società adottasse l'opinione che non vi siano differenze, e questa risultasse scorretta. Mi aspetterei che ne risultasse un male ben minore. Dati questi costi, io pretenderei delle prove che rendessero davvero molto improbabile l'ipotesi della similarità interrazziale, prima di ricusarla. La mia scelta scientifica dipende dai miei valori, non perché io sia acritico o perché mi piaccia credere che non vi sono simili differenze, ma perché l'unico modo di fare scelte coerenti nell'incertezza è considerare il costo degli errori alternativi che si possono fare. Al contrario, un ipotetico «scienziato neutro rispetto ai valori» presumibilmente accetterebbe di operare sulla base dell'ipotesi che tali differenze esistano, non appena vi fossero prove che la rendessero anche appena più probabile dell'ipotesi opposta.

Questi problemi non sorgono quotidianamente nel lavoro scientifico, perché i metodi statistici tradizionali di solito li includono nella più vasta categoria della scelta dei criteri di prova, o della tecnica particolare da usare per stimare qualche grandezza. Quella scelta viene poi fatta sul terreno convenzionale o tradizionale, usualmente senza discussione, giustificazione o perfino riconoscimento del fatto che si siano fatte scelte di valore [4].

Roberts ha scelto di illustrare il fatto che le ipotesi scientifiche non sono «libere da valori» citando quei valori che entrano nella scelta, da parte dello scienziato, di tollerare o non tollerare il rischio di sbagliare. I valori, come cercherò di dimostrare, entrano nelle scelte fatte dagli scienziati anche in altri (e, io credo, più importanti) modi. Per il momento, tuttavia, voglio solo affermare che è del tutto appropriato dire un «bravo!» ai biologi di cui abbiamo citato l'esempio, e dire «vergogna!» agli scienziati che recentemente hanno scritto che «un animale-macchina organizzato in simbiosi con un cervello e un apparato visivo animale per accrescere le funzioni meccaniche» sarà tecnicamente «fattibile» entro i prossimi quindici anni [5].

L'introduzione di parole come «etica» e «non si dovrebbe» nelle conversazioni che riguardano la scienza sembra sempre generare una tensione non dissimile, direi, dalla tensione che si sente salire ogni qual volta, durante una conversazione con anziani professori universitari tedeschi, capita di alludere alla carriera di uno dei loro colleghi che ebbe grande successo durante gli anni

[4] M.J. Roberts, *Nature and Condition of Social Science* («La natura e la situazione della scienza sociale»), in *Daedalus,* Estate 1974, pp. 54-55.

[5] O. Firschein, M.A. Fischler, L.S. Coles e J.M. Tenenbaum, *Intelligent Machines Are on the Way* («Sono in arrivo le macchine intelligenti»), *IEEE Spectrum*, luglio 1974, p. 43. L'opinione che ho citato qui sembra aver riscosso il consenso di « 41 esperti di intelligenza artificiale, compresi alcuni membri del Consiglio Internazionale Unito dell'intelligenza artificiale». L'articolo chiarisce esplicitamente che l'opinione è condivisa da tutti gli autori.

di Hitler. In quest'ultima situazione, il calo della temperatura sociale tradisce la paura che possa venir detto qualcosa di «spiacevole», e in particolare che possa venir notata la passata incapacità di quel collega di rinunciare alle sue ambizioni personali in ossequio alla moralità. Allora, naturalmente, ci si accorge che è messa in questione non solo la condotta di quel collega, bensì di tutti gli accademici tedeschi dell'epoca. Nell'altra situazione, la tensione tradisce una preoccupazione simile, perché l'etica, in fin dei conti, ha a che fare con qualcosa di più della rinuncia. La tensione tradisce la paura che venga detto qualcosa su ciò che la scienza, cioè gli scienziati, dovrebbe e non dovrebbe fare. E ci si accorge che quello di cui si potrebbe finire per parlare non riguarda solamente la scienza in generale, o una qualche popolazione astratta chiamata «scienziati», ma anche proprio le persone presenti.

Alcuni scienziati, anche se certamente non tutti, sostengono che il dominio della scienza è universale, che non ci può essere niente che, in conseguenza di un qualche principio «superiore», non dovrebbe essere studiato. E da questa premessa di solito si ricava la conclusione che qualsiasi discorso a base di «si dovrebbe», e applicato alla scienza, sia intrinsecamente sovversivo e anti-scientifico, anzi perfino anti-intellettuale.

Qualsiasi siano i pregi di questo ragionamento dal punto di vista della logica astratta, esso è idiota quando applicato a situazioni concrete, perché vi sono infiniti problemi aperti all'indagine scientifica, ma soltanto un numero finito di risorse al comando della scienza. L'uomo deve dunque scegliere quali problemi aggredire, e quali lasciare da parte. Noi non sappiamo, per esempio, se il numero dei pori sulla pelle di un individuo è in qualche modo correlato con il numero dei neuroni nel suo cervello. Per questo problema non vi è alcun interesse, e quindi nessuna controversia per stabilire se la scienza debba studiarlo. I cinesi hanno praticato l'agopuntura per secoli, senza sollevare l'interesse della scienza occidentale. Ora, improvvisamente, gli scienziati occidentali sono diventati interessatissimi. Questo esempio illustra il fatto che il «progresso»scientifico non si muove lungo un cammino determinato dalla natura stessa, e che esso invece riflette gli interessi e le preoccupazioni umane.

Sicuramente, l'intelligenza umana ben affilata è tra le risorse più scarse a disposizione della società moderna. E, chiaramente, alcuni dei problemi suscettibili di indagine scientifica sono più importanti di altri. La società umana deve perciò inevitabilmente affrontare il compito di distribuire in modo saggio quella risorsa scarsa che è il suo talento scientifico. C'è semplicemente una responsabilità — di cui non ci si può liberare con il desiderio — di decidere quali problemi sono più importanti, interessanti ecc. di altri. Ogni società deve continuamente trovare i modi di affrontare questa responsabilità. La domanda qui è *come*, in una società aperta, si debbano trovare questi modi; devono essere dettati, mettiamo, dall'*establishment* militare, o devono essere aperti al dibattito tra cittadini e scienziati? Se devono essere dibattuti, allora perché l'etica dovrebbe essere esclusa dalla discussione? E, infine, come è possibile che emerga qualcosa di sensato se, prima, non si è tutti d'accordo che, contrariamente a quanto sostenne John von Neumann, le possibilità tecnologiche non sono irresistibili per l'uomo? «Potere» non implica «dovere».

Purtroppo, il nuovo conformismo che ci permette di parlare di tutto, tranne che delle poche semplici verità che stanno scritte nel nostro cuore e nei libri sacri di ciascuna delle molte religioni dell'uomo, rende ogni argomentazione basata su queste verità — non importa quanto elegantemente pensata o quanto eloquentemente costruita — ridicola agli occhi degli scienziati e dei tecnici cui può essere indirizzata. Questo, di per sé, è forse l'esempio più tragico di come un'idea, se usata male, si trasformi nel suo opposto. Quegli scienziati che continuano a parlare a vanvera di «conoscenza fine a se stessa», per poter sfruttare quello slogan per i loro fini di utilità personale, hanno allontanato scienza e conoscenza da qualsiasi contatto col mondo reale. Una questione centrale per quanto riguarda la conoscenza, una volta che la si è guadagnata, è la sua convalida; ma quello che vediamo accadere in quasi tutti i campi, e specialmente nei rami dell'informatica di cui abbiamo discusso, è che la convalida della conoscenza scientifica è stata ridotta a esibizione di meraviglie tecnologiche. Questo si può interpretare soltanto in due modi: o la natura, cui la scienza si riferisce, consiste interamente in materia grezza da plasmare e manipolare come un oggetto, oppure la conoscenza che la ricerca scientifica ha fatto acquistare all'uomo è del tutto irrilevante per l'uomo stesso. La scienza non può ammettere che quest'ultima sia la realtà, perché se così fosse essa perderebbe la sua licenza di esercitare. Questa perdita, naturalmente, avrebbe conseguenze pratiche (coinvolgenti il denaro e tutto il resto) cui gli scienziati resisterebbero con tutta la loro forza. Se invece è vera la prima interpretazione, allora l'uomo stesso è diventato un oggetto. C'è abbondanza di prove che è questo, in realtà, ciò che è successo. Ma allora anche la conoscenza ha perso quella purezza di cui tanto si vantano gli scienziati; allora è diventata un'impresa né più né meno importante, né più intrinsecamente significativa, mettiamo, della conoscenza di come si organizza una catena di montaggio per automobili. Chi vorrebbe acquisire questa conoscenza come «fine a se stessa»?

Questa evoluzione è tragica, perché priva la scienza perfino della possibilità di essere guidata da standard autenticamente umani, mentre non limita in alcun modo la potenzialità della scienza nel conferire all'uomo una potenza sempre crescente. E anche qui troviamo le radici di quella disumanizzazione dell'uomo di cui tanto si parla. Un individuo viene disumanizzato ogni volta che viene trattato come qualcosa di meno di una persona nella sua interezza. Le varie forme di ingegneria umana e sociale che abbiamo discusso qui fanno proprio questo, poiché eludono tutti i contesti umani, e specialmente quelli che danno un significato reale al linguaggio umano.

Il fatto che le argomentazioni che fanno appello a principi superiori — mettiamo, ai doveri di un individuo nei confronti dei suoi figli, o alla natura stessa — non vengano riconosciute come legittime, crea un serio dilemma per chiunque desideri convincere i suoi colleghi a collaborare, imponendo certi limiti alla loro ricerca. Se egli sviluppa comunque argomentazioni simili, magari sperando di determinare una sorta di «esperienza di conversione» nei suoi colleghi, rischia allora di risultare completamente inefficace, e perfino di venire scomunicato più o meno come un buffone. Se sostiene l'idea delle limitazioni, sulla base del fatto che una ricerca non vincolata può causare conseguenze

irreversibili, allora partecipa anch'egli al rito della ragione strumentale (sotto forma, poniamo il caso, di analisi costi/benefici) contro cui intende lottare, e contribuisce a legittimarne l'abuso.

Così come accade per molti altri dilemmi, la soluzione di questo sta nel rifiuto delle regole del gioco che gli danno origine. Per questo dilemma, la regola operativa è che la salvezza del mondo — ed è proprio di questo che sto parlando — dipende dalla conversione degli altri alle idee giuste. Questa regola è falsa. La salvezza del mondo dipende soltanto dall'individuo di cui quello è il mondo. Per lo meno, ciascun individuo deve agire come se l'intero futuro del mondo, dell'umanità stessa, dipenda da lui. Qualsiasi cosa in meno di questo è una fuga dalle responsabilità, ed è di per sé una forza disumanizzante, perché incoraggia l'individuo a vedersi come soltanto un attore in un dramma scritto da sceneggiatori anonimi, come qualcosa di meno di una persona completa, e questo è l'inizio della passività e della demotivazione.

Questa non è una filippica in favore del solipsismo, né un consiglio dato a ogni uomo perché viva soltanto per se stesso. Ma sostiene, in effetti, che ogni uomo deve vivere *innanzitutto* per se stesso. Perché soltanto facendo esperienza del suo personale valore intrinseco, un valore affatto indipendente dall'«uso» che se ne può fare come strumento, egli potrà arrivare a conoscere quei fini che vanno al di là di lui e che in ultima analisi gli conferiscono la sua identità, e che sono i convalidatori ultimi della conoscenza umana.

Ma il fatto che ciascun individuo sia responsabile del mondo intero, e che l'adempimento di quella responsabilità coinvolga per prima cosa la responsabilità di ciascun individuo di fronte a se stesso, non implica che ciascuno di noi non abbia dei doveri nei confronti degli altri. Il primo tra questi è quello di istruirci l'un l'altro al meglio delle nostre possibilità. E la principale e più efficace forma di istruzione che possiamo mettere in atto è l'esempio che la nostra condotta fornisce a quanti ne vengono toccati. Gli insegnanti e gli scrittori hanno una responsabilità particolarmente pesante, proprio perché hanno raggiunto delle posizioni da cui il loro esempio è in grado di influenzare altre persone, oltre a quelle della loro cerchia immediata.

Lo spirito di queste considerazioni impone che io esponga alcune delle mie decisioni su quanto posso e non posso fare nel campo della scienza dei computer. Lo faccio con qualche timore, perché ho imparato che le persone si chiedono continuamente fra loro cosa debbano fare, mentre l'unica domanda davvero importante è cosa debbano essere. Il fisico Steven Weinberg, nel commentare certe recenti critiche alla scienza, scrive per esempio:

> Ho cercato di capire questi critici scorrendo alcuni dei loro scritti, e ho trovato parecchie considerazioni pertinenti, e anche toccanti. In modo particolare io condivido la loro diffidenza nei confronti di quanti, da David Ricardo al Club di Roma, applicano con troppa sicurezza i metodi delle scienze naturali alle faccende umane. Ma alla fine rimango interdetto. Che cos'è che vogliono che *io* faccia [6]?

[6] S. Weinberg, *Reflections of a Working Scientist* («Riflessioni di uno scienziato che lavora»), in *Daedalus*, estate 1974, p. 41.

Il mio timore è che si pensi che io stia rispondendo a una domanda del tipo di quella fatta da Weinberg. Non è questa la mia intenzione. Ma il rischio di poter essere frainteso non può esimermi dal mio dovere.

Non c'è, a mio parere, alcun progetto nella scienza dei computer in quanto tale che sia moralmente ripugnante, e che consiglierei a studenti e colleghi di evitare. I progetti di cui ho discusso, e altri che citerò, non fanno propriamente parte della scienza dei computer. I computer non sono un aspetto centrale del lavoro di Forrester e Skinner. E gli altri progetti citati non sono scienza dei computer, perché in massima parte non sono affatto scienza. Sono, come ho già suggerito, intelligenti aggregati di tecniche che puntano a far eseguire qualcosa. Forse a causa dell'accidente storico che ha fatto chiamare «scienza dei computer» quei dipartimenti universitari che si interessano di computer, tutto il lavoro fatto in quei dipartimenti viene indiscriminatamente chiamato «scienza», anche se solo una parte di esso merita l'onore di quell'appellativo. I rabberciatori tecnici (gli «adoratori di *gadget*», come li chiamava Norbert Wiener) talvolta trovano duro resistere alla tentazione di associarsi alla scienza e di sifonare un po' di legittimazione dal serbatoio che si è accumulato. Ma non tutti quelli che dicono di essere cantanti hanno una bella voce.

Non tutti i progetti che sono schiettamente orientati verso la prestazione sono pericolosi o moralmente ripugnanti. Molti di essi aiutano davvero l'uomo a portare avanti in modo più sicuro ed efficace il suo lavoro quotidiano. La navigazione e i meccanismi anti-collisione computerizzati, per esempio, permettono a navi e aerei di funzionare in situazioni che finora non erano affrontabili. La lista dei modi in cui il computer si è dimostrato utile è indubbiamente lunga. Ci sono, però, due tipi di applicazione del computer che non dovrebbero essere intrapresi affatto oppure, se vengono presi in considerazione, dovrebbero essere accostati con estrema cautela.

Quelle del primo tipo, le chiamerei semplicemente oscene. Sono quelle che, solo a prenderle in considerazione, dovrebbero dare un senso di disgusto a qualsiasi persona civile. La proposta che ho citato, di accoppiare un cervello e un sistema visivo a un computer, ne è un esempio. Rappresenta un attacco alla vita stessa. C'è da chiedersi cosa debba essere successo alla percezione della vita di chi l'ha proposta, e quindi alla sua percezione di sé come parte del continuo della vita, per fargli anche soltanto pensare una cosa simile, per non parlare del sostenerla in pubblico. A un livello molto più basso, c'è da chiedersi quale sia un bisogno umano concepibile che possa essere soddisfatto da un simile «apparecchio» unito ad altri, per non parlare di quell'apparecchio da solo.

Io metterei nella stessa categoria tutti i progetti che propongono di sostituire un sistema computerizzato a una funzione umana che coinvolge il rispetto interpersonale, la comprensione e l'amore. Perciò io rifiuto la proposta di Colby di installare computer psicoterapeutici, non perché un progetto simile sia infattibile dal punto di vista tecnico, ma perché è immorale. A giustificazione di quel progetto, ho sentito dire che una persona può ricevere un aiuto psicologico dalla conversazione con un computer, anche se si sa benissimo che il computer non «capisce» la persona. Un esempio che mi è stato citato era quello di un sistema computerizzato progettato per accettare testi in linguag-

gio naturale attraverso un terminale a tastiera, e per rispondere con una serie di «sì» e «no» generati a caso. Un paziente con problemi «conversava» con questo sistema e, a quanto si dice, ne veniva portato a pensare in modo più approfondito ai suoi guai, e ad arrivare a certe conclusioni, a quanto si dice, utili. Fino ad allora, quella persona era soltanto andata alla deriva, sommersa da preoccupazioni senza scopo. In linea di principio, un vassoio di biscotti della fortuna cinesi, o un mazzo di carte, avrebbero potuto avere lo stesso effetto. Il computer, però, ha aggiunto una certa aura — derivata, naturalmente, dalla scienza — che permetteva al «paziente» di crederci, laddove egli avrebbe respinto i biscotti della fortuna e le carte da gioco come strumenti di superstizione. Sorge allora la domanda (e si risponde da sé) se quello che vogliamo è incoraggiare la gente a condurre la sua vita sulla base della frode evidente, del ciarlatanismo, dell'irrealtà. E, cosa più importante, crediamo davvero che il preferire la terapia somministrata dalle macchine a quella condotta da altre persone possa aiutare la gente che vive nel nostro mondo, già apertamente strutturato come una macchina? Ho sentito rispondere a quest'ultima domanda con l'affermazione secondo cui la mia posizione non è niente di più che un «... o mangi questa minestra...». Dicono che tale posizione ignora la carenza di buoni psicoterapeuti umani, e che nega alle persone con problemi quel po' d'aiuto che i computer possono oggi dar loro, soltanto perché i computer attualmente disponibili non sono «ancora» all'altezza, mettiamo, della migliore psicoanalisi. Ma quell'obiezione manca completamente il punto essenziale. Il punto è (nonostante Simon e Colby siano contrari) che ci sono funzioni umane cui non si *deve* sostituire il computer. Non ha niente a che vedere con quello che si può, o meno, riuscire a far fare ai computer. Rispetto, comprensione e amore non sono problemi tecnici.

Il secondo tipo di applicazione del computer che andrebbe evitato, o almeno non andrebbe intrapreso senza pensarci prima molto bene, è quello di cui si vede facilmente che ha effetti collaterali irreversibili e non del tutto prevedibili. Se, oltre a questo, una tale applicazione non risponde palesemente a un bisogno umano che non può essere facilmente soddisfatto altrimenti, allora non la si dovrebbe portare avanti. Quest'ultima restrizione consegue direttamente dall'argomentazione, che ho già presentato, sulla scarsità dell'intelligenza umana.

L'esempio che voglio citare a questo punto è il riconoscimento automatico della voce umana. Ci sono oggi tre o quattro progetti, negli Stati Uniti, dedicati a mettere in grado dei computer di capire la voce umana, cioè a programmarli in modo tale che un discorso umano loro rivolto possa essere convertito in quella stessa rappresentazione interna che risulterebbe se quello che si dice venisse invece inserito nel computer usando una tastiera.

Il problema, come si può ben vedere, è molto più complicato di quello della comprensione del linguaggio naturale in quanto tale, perché per poter seguire il flusso di un discorso coerente, in primo luogo bisogna che venga capito il linguaggio in cui quel discorso viene fatto. La soluzione del «problema della comprensione della voce» presuppone perciò la soluzione del «problema della comprensione del linguaggio naturale». E abbiamo visto che, per que-

st'ultimo, possediamo soltanto «una parte minuscola delle conoscenze necessarie». Ma qui non mi preoccupo della fattibilità tecnica dell'impresa, né di stimare quanto o quanto poco ottimisti si possa essere sul suo completamento.

Perché dovremmo sobbarcarci quest'impresa, tanto per cominciare? Ho fatto questa domanda a molti entusiasti del progetto. La risposta più pronta che sono riuscito ad avere è che si aiuteranno i dottori a registrare le loro osservazioni mediche, e quindi a tradurle in azione in modo più efficace. Naturalmente, tutto ciò che ha un qualsiasi legame con la medicina viene automaticamente considerato buono. Ma qui dobbiamo ricordare che il problema è così enorme che soltanto i più grandi computer riusciranno mai a trattarlo. In altre parole, se anche il sistema desiderato venisse progettato con successo, richiederebbe con tutta probabilità un computer così grande, e quindi così costoso, che soltanto gli ospedali più grandi e facoltosi potrebbero permetterselo; ma in realtà l'intero sistema potrebbe essere così proibitivamente caro che neanche quelli potrebbero permetterselo. La domanda diviene allora: è proprio questo ciò di cui oggi la medicina ha più bisogno? Non sarebbe meglio impiegare tutto quel talento, per non parlare del denaro e delle risorse che esso rappresenta, per progetti che affrontino problemi più urgenti e fondamentali nel campo della cura della salute?

Ma poi, questa sedicente giustificazione della «ricerca» sul riconoscimento della voce è comunque soltanto una razionalizzazione (metto la parola «ricerca» tra virgolette, perché il lavoro di cui sto discutendo qui è di puro rappezzo. Non ho nessuna obiezione da fare agli scienziati seri che studiano la psico-fisiologia del riconoscimento del discorso umano). Se domande di quel tipo vengono fatte al principale *sponsor* di questo lavoro, l'Agenzia per i progetti di ricerca avanzata (ARPA) del Dipartimento della Difesa degli Stati Uniti, come è accaduto di recente a una conferenza aperta, la risposta che viene data è che la Marina spera di controllare le sue navi, e gli altri servizi legati alle sue armi, con comandi vocali. Questo progetto rappresenta allora, agli occhi del suo principale finanziatore, un lungo passo nella direzione di un campo di battaglia completamente automatizzato. Non vedo alcuna ragione per consigliare ai miei studenti di prestare i loro talenti a quello scopo.

Ho invitato i miei studenti e i miei colleghi a porre ancora un'altra domanda su questo progetto: posto che una macchina per il riconoscimento della voce non può che essere enormemente costosa, e che quindi soltanto governi e pochissime industrie molto grandi potranno permettersela, per che cosa la useranno? Per che cosa potrà mai essere usata? Per me, è fuori discussione che non c'è alcun problema umano urgente che verrà risolto più facilmente grazie all'esistenza di una macchina simile. Ma queste macchine ascoltatrici, se potessero essere realizzate, renderebbero la sorveglianza delle comunicazioni verbali molto più facile di quanto sia oggi. Forse l'unica ragione per cui, in molti Paesi del mondo, il governo controlla in modo molto limitato le conversazioni telefoniche tra i cittadini, è che quel controllo richiede moltissimo personale. Ciascuna conversazione fatta usando un telefono controllato deve prima o poi venire ascoltata da un agente umano. Ma le macchine per il riconoscimento della voce potrebbero distruggere tutte le conversazioni «non inte-

ressanti», e produrre trascrizioni soltanto delle altre ai loro padroni. Non credo neanche per sogno che raggiungeremo la capacità di fare questo in quel futuro prossimo tanto chiaramente visibile a Newell e Simon. Ma io chiedo, perché mai un tecnologo dei computer dotato di talento dovrebbe offrire il suo apporto a un progetto simile? Come cittadino mi chiedo, perché il mio governo spende circa due milioni e mezzo di dollari ogni anno (come fa attualmente) su questo progetto?

Certamente, domande di questo tipo si sono presentate a persone ponderate anche nelle fasi precedenti di scienza e tecnologia. Ma fino a poco tempo fa la società poteva sempre affrontare gli effetti indesiderati e pericolosi delle sue nuove invenzioni riorganizzandosi, in un certo senso, per disfare o minimizzare quegli effetti. La densità di popolazione nelle città poteva venire ridotta estendendo geograficamente le città. Un individuo poteva evitare i terribili effetti della rivoluzione industriale in Inghilterra, trasferendosi nel Nord America. E l'America poteva sfuggire a molte delle conseguenze della crescente potenza degli ordigni militari, ritirandosi dietro le sue fortificazioni oceaniche. Ma quei giorni sono passati. Lo scienziato e il tecnologo non possono più eludere la responsabilità per quello che fanno, richiamandosi all'infinita capacità della società di trasformarsi in risposta a nuove realtà, e di sanare le ferite che essi le infliggono. Si sono raggiunti certi limiti. Le trasformazioni che le nuove tecnologie possono richiedere possono essere impossibili da raggiungere, e il loro mancato raggiungimento può significare l'annientamento dell'intera vita. Nessuno ha il diritto di imporre all'umanità una scelta come questa.

Ho parlato qui di cosa si dovrebbe e non si dovrebbe fare, di cosa è moralmente ripugnante, e di cosa è pericoloso. Naturalmente, so bene che questi miei giudizi non hanno di per sé alcuna forza morale, se non su me stesso. Né, come ho già detto, ho alcuna intenzione di dire agli altri quali compiti dovrebbero e non dovrebbero intraprendere. Li invito solo a considerare le conseguenze di quello che fanno. E qui non intendo soltanto, e nemmeno principalmente, le conseguenze dirette delle loro azioni sul mondo che li circonda. Intendo piuttosto le conseguenze su loro stessi, quando costruiscono le loro razionalizzazioni, quando reprimono le verità che li spingono in direzioni diverse, e quando fanno a pezzi la loro stessa autonomia. Il fatto che tanta gente chieda così spesso cosa deve fare è un segno del fatto che l'ordine di essere e fare si è invertito. Coloro che sanno chi e che cosa sono, non hanno bisogno di chiedere cosa debbano fare. E coloro che sono costretti a chiederlo non potranno smettere di farlo finché non cominceranno a guardare dentro se stessi. Ma è compito di ognuno mostrare con l'esempio quali domande si possono fare a se stessi, e mostrare che si può vivere con le poche risposte che ci sono.

Ma proprio come io non ho il diritto di stabilire le azioni degli altri, così i costruttori del mondo in cui devo vivere non hanno il diritto di imporre senza condizioni le loro visioni a me. Scienziati e tecnologi hanno, a causa del loro potere, una responsabilità particolarmente gravosa, che non deve essere schivata rifugiandosi dietro una facciata di slogan come quello dell'ineluttabilità tecnologica. In un mondo in cui l'uomo sempre più incontra solo se stesso, e solo sotto forma dei prodotti che ha creato, chi fa e progetta questi prodotti

— edifici, aeroplani, generi alimentari, bombe e così via — deve essere profondamente consapevole del fatto che i suoi prodotti sono, in fin dei conti, il risultato di scelte umane. Gli uomini potrebbero invece scegliere di avere automobili veramente sicure, televisori decenti, una casa decente per tutti, o mezzi di trasporto collettivo sicuri, comodi e largamente distribuiti. Il fatto che queste cose non esistano, in un Paese che ha le risorse per produrle, è conseguenza non dell'ineluttabilità tecnologica, bensì del fatto che la gente ha scelto di produrre e avere esattamente le cose che abbiamo prodotto e che in realtà abbiamo.

È duro, quando si assiste a un annuncio pubblicitario televisivo particolarmente disgustoso, immaginare che degli esseri umani adulti, un certo giorno in un certo posto, si siano seduti intorno a un tavolo per creare esattamente quell'annuncio e farlo trasmettere centinaia di volte. Ma questo è ciò che succede. Queste cose non sono il prodotto di forze anonime. Sono il prodotto di gruppi di uomini che si sono accordati fra loro, decidendo che questo inquinamento della coscienza della gente serve ai loro scopi.

Ma, come è stato fin dall'inizio della storia documentata, le decisioni che hanno le conseguenze più nefaste vengono spesso prese al servizio di qualche buon fine primario. Per esempio, nell'estate del 1966 c'era negli Stati Uniti una notevole inquietudine a proposito dei bombardamenti intensivi americani nel Vietnam del Nord (la distruzione seminata dai bombardieri americani nel Vietnam del Sud non era altrettanto al centro del dibattito pubblico, perché l'opinione pubblica era ancora convinta che gli Stati Uniti stessero «aiutando» quel Paese disgraziato). Una quarantina di scienziati di alto livello nella gerarchia scientifica decise di contribuire a fermare i bombardamenti, indicendo un gruppo di studio estivo sotto gli auspici dell'*Institute of Defense Analyses,* una prestigiosa ditta consulente del Dipartimento della Difesa. Intendevano dimostrare che il bombardamento era in realtà inefficace [7].

Fecero la loro dimostrazione usando i migliori strumenti scientifici, ricerca operativa, analisi dei sistemi e tutto il resto. Ma ebbero l'impressione che non sarebbero stati ascoltati dal Ministero della Difesa, a meno che suggerissero un'alternativa al bombardamento. Proposero di piazzare uno «steccato elettronico» nella zona cosidetta smilitarizzata che separava il Vietnam del Nord da quello del Sud. Questa barriera doveva fermare gli infiltratori provenienti dal Nord. Doveva consistere, tra l'altro, in piccole mine sepolte nel terreno e progettate per esplodere se calpestate dai piedi dei portatori, ma per essere insensibili al passaggio degli autocarri. Altri apparecchi dovevano interdire il traffico di autocarri. I vari sensori elettronici, i loro sistemi di controllo ecc. finirono per diventare parte della cosidetta «linea McNamara». Questo fu l'inizio di quello che da allora è evoluto nel concetto di «campo di battaglia elettronico».

L'intento della maggior parte di questi uomini non era inventare o raccomandare una nuova tecnologia che avrebbe reso la guerra più terribile e, per

[7] Vedere l'edizione Gravel di *The Pentagon Papers* («Le carte del Pentagono»), Beacon Press, Boston, Mass. 1971, volume IV, specialmente p. 115.

inciso, meno costosa per le nazioni altamente industrializzate, a spese di quelle «sottosviluppate». Il loro intento era fermare i bombardamenti. In questo, essi erano completamente dalla parte dei gruppi pacifisti, e dei cittadini di buon senso in generale. E in effetti raggiunsero il loro obiettivo; per un po', il bombardamento del Vietnam del Nord venne cessato, e fu installata la linea McNamara. Tuttavia, queste persone di enorme impatto e influenza avrebbero potuto invece semplicemente annunciare che credevano che i bombardamenti, anzi l'intera avventura americana in Vietnam, fossero sbagliati, e che non avrebbero più «contribuito». Io so che almeno alcuni tra i partecipanti pensavano che la guerra fosse sbagliata; forse lo pensavano tutti. Ma, come alcuni di loro mi spiegarono in seguito, sentivano che se avessero fatto un simile annuncio non sarebbero stati ascoltati, né allora né mai. Eppure, chi può dire quale sarebbe stato l'effetto se quaranta tra i principali scienziati americani, nell'estate del 1966, si fossero uniti ai gruppi pacifisti nel condannare semplicemente la guerra per ragioni morali? E a parte l'effetto positivo che una simile presa di posizione avrebbe potuto avere sugli avvenimenti mondiali, quale è stato l'effetto negativo del loro compromesso su loro stessi, sui loro colleghi e sugli studenti cui essi erano di esempio?

Sono parecchie le lezioni da trarre da questo episodio. La prima è che non è stata l'ineluttabilità tecnologica a inventare il campo di battaglia elettronico, e neppure un insieme di forze anonime. Degli uomini, proprio come quelli che progettano gli annunci pubblicitari televisivi, si sono seduti intorno a un tavolo e hanno scelto. Eppure, il risultato dei dibattiti dello Studio Estivo del 1966 erano in un certo senso preordinati. Il campo delle risposte che si ottengono è determinato dal dominio delle domande che si fanno. Non appena fu stabilito che lo Studio Estivo doveva interessarsi soltanto di questioni tecniche, la soluzione al problema di fermare i bombardamenti nel Vietnam del Nord divenne sostanzialmente una questione di calcoli. Quando si aggiunse l'ulteriore condizione che il gruppo dovesse a tutti i costi conservare la sua credibilità con i suoi *sponsor*, che non dovesse mettere in pericolo lo status di «persone interne al sistema» dei partecipanti, allora tutti i gradi di libertà che i suoi membri potevano aver avuto inizialmente andarono in pratica perduti. Molti dei partecipanti, lo so, hanno difeso la libertà accademica, la loro come quella di colleghi la cui carriera era in pericolo per ragioni politiche. Questi uomini non sentirono di rischiare le loro libertà accademiche o di studiosi, quando si impegnarono nel tipo di consulenza proprio dello Studio Estivo. Ma il sacrificio dei gradi di libertà che avrebbero potuto avere se non si fossero slanciati in modo così totale verso le esigenze dei loro finanziatori (che questo sacrificio sia stato fatto volontariamente o meno) è stato una forma di censura più poderosa di quella che avrebbe potuto essere imposta da ufficiali governativi. Questa sorta di auto-mutilazione intellettuale, proprio perché è in larga misura inconscia, è una delle fonti principali del senso di impotenza provato da tante persone che, almeno superficialmente, sembrano occupare posizioni di potere.

Una seconda lezione è questa. Questi uomini poterono dare la consulenza che diedero perché operavano a una distanza psicologica enorme dalla gente che sarebbe stata mutilata e uccisa dai sistemi d'arma che sarebbero risultati

dalle idee comunicate da loro ai loro patrocinatori. La lezione perciò, è che lo scienziato e il tecnologo devono, con atti di volontà e di fantasia, impegnarsi attivamente per ridurre queste distanze psicologiche, per opporsi alle forze che tendono ad allontanarli dalle conseguenze delle loro azioni. Essi devono pensare a quello che stanno realmente facendo, nient'altro che questo. Devono imparare ad ascoltare la loro voce interna. Devono imparare a dire «No!».

In ultima analisi, è l'azione stessa che importa. Quando la ragione strumentale è l'unica guida per l'azione, gli atti che essa giustifica sono derubati del loro significato intrinseco, ed esistono così in un vuoto etico. Recentemente ho sentito un dirigente di una grande università difendere pubblicamente un'importante decisione politica da lui presa, e osteggiata per ragioni morali da molti degli studenti e degli insegnanti di quell'università, con le parole: «Avremmo potuto prendere una posizione morale, ma a che cosa sarebbe servito?». Ma la bontà di un atto morale è una proprietà intrinseca dell'atto stesso. È per questo che un atto può di per sé nobilitare o corrompere la persona che lo effettua. La vittoria della ragione strumentale nel nostro tempo ha determinato in pratica la scomparsa di questa consapevolezza, e così a maggior ragione la de-legittimazione dell'idea stessa di nobiltà.

Sono consapevole, naturalmente, che quasi nessuno di coloro che leggono queste righe avrà l'impressione di esserne riguardato; tanto è profondo il modo in cui la convinzione che siamo tutti dominati da forze anonime, al di là del nostro controllo, è penetrato nella coscienza generale del nostro tempo. E questa convinzione è accompagnata da uno svilimento dell'idea di coraggio civile.

È una credenza diffusa, ma tristemente erronea, quella per cui il coraggio civile trova modo di esercitarsi soltanto nel contesto di avvenimenti che scuotono il mondo. Al contrario, il suo esercizio più arduo ha spesso luogo in quei piccoli contesti in cui la sfida è quella di superare i timori indotti da futili preoccupazioni di carriera, delle nostre relazioni con coloro che sembrano aver potere su di noi, o di qualsiasi cosa che possa turbare la tranquillità della nostra esistenza quotidiana.

Se questo libro deve essere visto come il difensore di qualcosa, allora facciamo che sia un richiamo a questo semplice tipo di coraggio. E poiché, dopo tutto, questo libro è sui computer, quel richiamo venga ascoltato principalmente dagli insegnanti di scienza dei computer.

Voglio che mi abbiano sentito affermare che il computer è una nuova, potente metafora per aiutarci a capire molti aspetti del mondo, ma che schiavizza quella mente che non abbia altre metafore, e abbia poche altre risorse, cui fare ricorso.

Il mondo è molte cose, e nessuno schema singolo è grande abbastanza da contenerle tutte, né quello della scienza dell'uomo né quello della sua poesia, né quello della ragione calcolatrice né quello della pura intuizione. E proprio come l'amore per la musica non basta per saper suonare il violino — bisogna anche saper padroneggiare l'arte dello strumento, e della musica stessa — così l'amore per l'umanità non basta per aiutarla a sopravvivere. La vocazione dell'insegnante a insegnare la sua arte è quindi degna di rispetto. Ma egli

deve fare di più: deve insegnare più di una metafora, e deve insegnare più con l'esempio della sua condotta che con ciò che scrive sulla lavagna. Deve insegnare i limiti dei suoi strumenti, oltre che la loro potenza.

Per combinazione, programmare è un'arte che si impara facilmente. Quasi chiunque abbia capacità mentali ragionevolmente normali può diventare un programmatore piuttosto bravo, con soltanto un po' di spiegazioni e di esercizio. E siccome la programmazione gratifica quasi immediatamente, vale a dire, siccome un computer inizia molto rapidamente a comportarsi più o meno nel modo voluto dal programmatore, la programmazione è molto seducente, specialmente per i principianti. Per di più, attrae maggiormente proprio coloro che non hanno ancora una maturità sufficiente per sopportare lunghi ritardi tra uno sforzo per ottenere qualcosa e la comparsa di prove concrete del successo. Gli studenti immaturi vengono perciò facilmente indotti a pensare di essere giunti a padroneggiare un'arte di immensa potenza e grande importanza quando, in realtà, ne hanno imparato appena i rudimenti, e proprio niente di sostanziale. La rapida ascesa di uno studente, da uno stato di completa ignoranza sui computer a quella che sembra essere maestria nella programmazione (ma in realtà è solo un livello costante piuttosto basso) può lasciare in lui un euforico senso di conquista, e una convinzione di aver trovato la sua vera vocazione. L'insegnante, naturalmente, tende anch'egli a sentirsi gratificato dall'ovvio entusiasmo di questi studenti, e quindi a incoraggiarlo, magari inconsciamente e contro il suo stesso giudizio più ponderato. Ma per lo studente questa può benissimo essere una trappola. Può impegnarsi in modo così totale in quella che ingenuamente crede essere la scienza dei computer, vale a dire nel semplice raffinamento delle sue capacità di programmazione, da precludersi in effetti lo studio di qualcosa di sostanziale.

Purtroppo, molte università hanno programmi di «scienza del computer» durante il corso di laurea, che permettono agli studenti di seguire questo indirizzo, anzi li incoraggiano a farlo. Quando questi studenti hanno completato i loro studi, essi sono in una situazione simile a quella di una persona che abbia imparato a parlare in una lingua straniera ma che, quando cerca di scrivere qualcosa in quella lingua, scopre di non avere letteralmente niente da dire.

La lezione qui è che, benché l'apprendimento di un'arte sia importante, esso non può essere tutto.

La funzione di un'università non può essere semplicemente quella di offrire ai potenziali studenti un catalogo di «esperienze» da cui scegliere. Perché se fosse quella la sua funzione, allora l'università dovrebbe supporre che gli studenti che arrivano siano già diventati quello che devono diventare. L'università avrebbe allora sicuramente ragione nel considerare lo studente come un cestino da supermercato, da riempire di generi diversi scelti nell'inventario intellettuale dell'università. Avrebbe ragione, in altre parole, nel vedere lo studente come un oggetto molto simile a un computer, le cui banche di memoria sono sempre affamate di nuovi «dati». Ma di sicuro questa non può essere una caratterizzazione adeguata di ciò che un'università è o dovrebbe essere. Di sicuro, l'università dovrebbe guardare a ciascun suo cittadino, studenti e docenti sullo stesso piano, prima di tutto come a un essere umano in cerca di verità

(come chiamarla altrimenti?), e quindi in cerca di se stesso. A ciascun membro dell'università dovrebbe continuamente accadere qualcosa di significativo; ognuno dovrebbe lasciare le sue aule essendo diventato un'altra persona, rispetto a quella che ci era entrata al mattino. Il semplice insegnamento di un mestiere non può adempiere a questa alta funzione dell'università.

Proprio perché tanta parte di un corso di laurea in scienza dei computer ha a che fare con l'arte del calcolo, è forse facile per l'insegnante scivolare nell'abitudine di dare soltanto un addestramento tecnico. Ma, se lo farà, sicuramente sminuirà se stesso e la sua professione. Si distaccherà anche dal resto della vita intellettuale e morale dell'università. L'università dovrebbe tener ferma, di fronte a ciascuno dei suoi cittadini, e anche di fronte al mondo più generale, una visione di cosa un uomo o una donna possono diventare. Fa questo dando sempre nuova vita alle idee degli uomini e delle donne che, in virtù delle loro conquiste, hanno contribuito a costruire la casa in cui viviamo. E lo fa, nel bene o nel male, per mezzo dell'esempio che ciascuno dei cittadini dell'università è per ciascun altro. L'insegnante di scienza dei computer, né più né meno di ogni altro membro del corpo docente, invita in pratica continuamente i suoi studenti a diventare quello che egli stesso è. Se vede se stesso come un semplice allenatore, un semplice applicatore di «metodi» per raggiungere fini determinati da altri, allora fa ai suoi studenti un doppio disservizio. Per prima cosa, li invita a diventare semplici esecutori degli ordini di altri, in ultima analisi nulla di meglio di quelle macchine che potrebbero un giorno sostituirli in quella funzione. In secondo luogo, impedisce loro di dare un'occhiata a quelle idee che, sole, fanno meritare alla scienza dei computer un posto nei programmi universitari. E nel fare ciò, chiude i loro occhi davanti agli esempi che scienziati del computer che si sono comportati come esseri umani creativi possono aver fornito, e quindi li priva della loro migliore *chance* di diventare essi stessi dei bravi scienziati del computer [8].

Infine, l'insegnante di informatica è soggetto egli stesso all'enorme tentazione di essere arrogante, perché la sua specialità è in qualche modo «più difficile» di quella dei colleghi umanisti. Ma la difficoltà delle nozioni a sua disposizione non è di alcun vantaggio. Le sue conoscenze sono semplicemente meno ambigue e perciò, come i suoi linguaggi per il computer, meno espressive della realtà. Gli umanisti in particolare

> hanno una maggiore familiarità con un mondo (morale) ambiguo, intrattabile, talvolta irraggiungibile, che non si ridurrà ad alcuna corrispondenza con i simboli per mezzo dei quali si potrebbe cercare di misurarlo. C'è un mondo che si tiene alla larga da tutti gli sforzi degli storici per ridurlo alle leggi della storia,

[8] Quasi immediatamente dopo aver scritto questi ultimi capoversi, ho ricevuto un articolo intitolato *Methodology and the Study of Religion, some Misgivings* («La metodologia e lo studio della religione; alcuni timori») da W.C. Smith, professore di religione alla Dalhousie University di Halifax, Nova Scotia. Il suo articolo esprime, oltre a molte altre idee importanti, molti dei punti che ho citato qui. Dal momento che ho letto alcuni dei precedenti lavori del professor Smith, non posso fare a meno di pensare che gli devo alcune delle idee espresse qui.

> un mondo che vanifica tutti gli sforzi degli artisti per capire le sue leggi fondamentali di bellezza. Lo sforzo [...] dell'uomo dovrebbe essere dedicato a una conoscenza più «morbida» di quella scientifica. [...] Questa non è una ritirata, bensì un passo avanti [9].

L'insegnante di scienza dei computer deve avere il coraggio di resistere alla tentazione dell'arroganza e di insegnare, di nuovo per lo più con il suo esempio personale, la validità e la legittimità delle conoscenze più morbide. Perché parlare di coraggio a questo proposito? Per due ragioni. La prima, e la meno importante, è che quanto più egli riesce a insegnare in questo modo, tanto più rischia la censura dei colleghi che, avendo meno coraggio di lui, hanno finito per soccombere alla visione semplicistica del mondo tipica del garantire diritti imperiali alla scienza. La seconda è che, se egli deve insegnare queste cose con il suo esempio personale, deve avere il coraggio di riconoscere, per usare le parole di Jerome Bruner, i prodotti della sua soggettività.

In precedenza ho paragonato l'inconscio a un mare turbolento, e la frontiera che divide la mente conscia, logica, dall'inconscio, a una linea costiera tempestosa. Quell'analogia è utile anche qui. Perché il coraggio che ci vuole per esplorare una costa pericolosa è simile a quello che bisogna mettere insieme per riuscire a sondare il proprio inconscio, per accogliere nel cuore e nella mente ciò che esso fa arenare sulla spiaggia della coscienza, e per esaminarlo nonostante la propria paura. Perché l'inconscio fa arenare non soltanto il materiale della creatività, non soltanto perle che vanno soltanto smerigliate, prima di fissarle a strutture di cui poi si può parlare con orgoglio, ma anche le più oscure verità sul nostro io. Anche queste vanno esaminate, comprese e in qualche modo incorporate nella nostra vita.

Se l'insegnante (se chiunque) deve essere per gli altri un esempio di una persona nella sua interezza, egli dovrà prima sforzarsi di essere una persona completa. Senza il coraggio di affrontare il nostro mondo interiore, oltre che quello esterno, questa interezza è impossibile da raggiungere. La ragione strumentale, da sola, non può portare a essa. E proprio lì sta una differenza cruciale tra uomo e macchina: l'uomo, per poter diventare completo, deve continuamente e per sempre essere un esploratore delle sue realtà, tanto interiori che esterne. La sua vita è piena di rischi, ma rischi che egli ha il coraggio di accettare perché, come l'esploratore, egli impara a fare affidamento sulle sue capacità di resistenza, di superamento delle difficoltà. Che significato potrebbe avere parlare di rischio, coraggio, affidamento, resistenza, e superamento delle difficoltà, quando si parla di macchine?

[9] C. Oglesby, *A Juanist Way of Knowledge* («Un modo di conoscere juanista»), lezione presentata al Seminario di tecnologia e cultura del MIT, il 26 ottobre 1971. Copie si possono avere dal Rev. John Crocker jr., 312 Memorial Drive, Cambridge MA 02139, USA.

APPENDICE

NON SENZA DI NOI

(traduzione di un discorso tenuto da Joseph Weizenbaum alla Società di Informatica di Karlsruhe, il 17 luglio 1986)

Ogni volta che vengo in Europa, e specialmente in Germania Federale, sono sorpreso dalla normalità della vita quotidiana. Come visitatore occasionale della Germania e del continente europeo, vedo cose e osservo circostanze che a voi devono sembrare di routine, o addirittura naturali. Per esempio, le vostre super-autostrade, la «musica» che ti assale nei ristoranti, gli ascensori, i molti parcheggi, le foreste di antenne televisive che popolano i tetti delle vostre case, e così via. Tutti questi fenomeni sono probabilmente così usuali che non ve ne accorgete più. Anche negli Stati Uniti abbiamo tutte queste cose, ed esse sono altrettanto invisibili a noi.

Ma a dispetto della profonda americanizzazione delle vostre città e della vostra campagna, qui rimangono alcune cose che sono europee e non possono venire americanizzate, venir fatte sembrare normali o rese invisibili. Almeno, così mi sembra. O forse m'inganno? Mi riferisco alla vostra posizione geografica, e a tutto quello che ne segue. Mi riferisco, per esempio, al vostro confine con l'altra Germania, pieno di installazioni militari di ogni sorta. Mi riferisco a quei tombini delle vostre strade destinati a essere riempiti di mine nucleari, e mi riferisco alle piccole distanze che vi separano da tutti i magazzini di armi nucleari, e così via. In altre parole, alla vostra prossimità puramente fisica, ma anche psicologica, alla catastrofe finale.

Negli Stati Uniti noi non siamo, in un certo senso, più distanti dalla catastrofe di quanto lo siate voi qui. Non soltanto Chernobyl, ma anche la minaccia della guerra è ovunque. E la guerra è il nemico di tutti! In caso di guerra, non importa se iniziata in modo non intenzionale da una tecnologia ufficialmente destinata ad evitare la guerra, o da cosiddetti uomini (o donne) di Stato che hanno creduto fosse loro dovere schiacciare il bottone, voi forse morirete dieci minuti prima di noi che stiamo nella fortezza-America, ma tutti moriremo. Noi, però, non abbiamo nelle strade tombini per le mine nucleari destinate a ritardare l'avanzata dei càrri armati sovietici. Noi vediamo i silos dei nostri missili solo di tanto in tanto, vale a dire quando qualcuno decide di mostrarceli in televisione. Per quanto impegno il nostro governo metta nel convincerci che i malvagi russi sono in realtà vicini a noi come a voi europei, per esempio che ci minacciano da Cuba e dal Nicaragua, gli americani nel com-

plesso restano non convinti, e perciò non turbati, da quegli sforzi (dopo tutto, siamo riusciti all'ultimo minuto a domare Grenada, grazie ai molti miliardi di dollari che ci hanno permesso di raggiungere un livello di armamento paragonabile alla potenza militare di Grenada. Grazie, Ronald Reagan!).

D'altra parte, se l'americano medio fosse consapevole del pericolo che tutti abbiamo di fronte, questo sarebbe forse più stupefacente del fatto che se ne preoccupi così poco. L'esperienza che gli Stati Uniti hanno della guerra permette il crescere di una mentalità del tipo «qui non può succedere», piuttosto che di un timore concreto di qualcosa che sembra davvero essere lontanissimo dalle preoccupazioni immediate della vita quotidiana.

Mi rendo anche conto del fatto che è emotivamente impossibile per voi — come per chiunque — vivere molto a lungo di fronte a minacce alla vostra stessa esistenza, senza mettere in azione dei meccanismi psicologici che vi permettano di escludere questi pericoli dalla vostra coscienza, concedendo loro, al massimo, di emergere ogni tanto. Ma quando una simile repressione impone sforzi sistematicamente male indirizzati, o esclude un comportamento potenzialmente salvifico, allora è il momento di sostituirla con un'analisi profonda dell'abisso stesso.

Questo momento è venuto, per i professionisti del computer. Ora noi abbiamo il potere di cambiare in modo fondamentale lo stato del mondo, e in direzioni che conducono alla vita. Per trovare il coraggio necessario — non tutti siamo santi o eroi — dobbiamo capire che per noi come individui, e per coloro che ci sono cari, il nostro comportamento attuale è molto più pericoloso, anzi di minaccia alla vita stessa, di quello che ci richiederebbe un sano buon senso.

Lo dirò in modo molto semplice: è una verità banale che nessuna delle armi e dei sistemi d'arma che oggi minacciano di sterminio ogni essere umano, e il cui progetto, produzione e vendita minacciano di portare miseria a tutti i popoli della terra e condannano innumerevoli persone, e specialmente bambini, alla morte per fame, nessuno di questi ordigni avrebbe potuto essere sviluppato senza la zelante, e perfino entusiastica, collaborazione dei professionisti del computer. Non può andare avanti senza di noi! Senza di noi la corsa agli armamenti, e specialmente quella qualitativa, non può avanzare di un altro passo.

Ci dice niente questo fatto evidente, semplice e ovvio? Io penso di sì. Per prima cosa, ci dice che gli esperti del computer — come anche gli specialisti in molti altri campi tecnici — condividono la responsabilità di aver determinato l'attuale, pericoloso stato del mondo. Questo dovrebbe rendere chiaro che quelli tra noi che, forse senza rendersene conto, dedicano il loro talento e le loro forze alla morte piuttosto che alla vita, non hanno diritto di dare ai politici, agli uomini e alle donne di Stato, la colpa per non saperci portare alla pace. Senza di noi, essi non potrebbero mettersi in pericolo tra loro, né mettere in pericolo i popoli della nostra Terra. Ma non basta fare dei bei cartelloni da portare in imponenti manifestazioni. Coloro che li portano devono chiedersi se il loro lavoro quotidiano non contribuisca a rendere possibili proprio quegli ordigni contro cui stanno protestando.

A questo proposito, viene soprattutto in mente il settore chiamato «Intelligenza Artificiale». Molti dei problemi tecnici che questa sotto-disciplina della scienza dei computer si trova ad affrontare stimolano con particolare forza la fantasia e la creatività dei ricercatori a orientamento tecnico. Obiettivi come fare di un computer un essere pensante, dare al computer la capacità di capire il linguaggio parlato, mettere il computer in grado di vedere, obiettivi come questi offrono tentazioni quasi irresistibili a quelli fra noi che non hanno del tutto sublimato la tendenza a giocare con paletta e secchiello, o che intendono trovare sul palcoscenico del computer (cioè usando sistemi di computer) soddisfazione alle loro illusioni di onnipotenza. Imprese come queste sono straordinariamente impegnative e interessanti. Robert Oppenheimer le chiamava «dolci». Inoltre, i progetti di ricerca in questi settori vengono finanziati con generosità. Il denaro necessario esce di solito dai forzieri dei militari, almeno negli Stati Uniti.

È enormemente attraente e, nel lavoro dell'Intelligenza Artificiale, seducentemente semplice perdersi o nascondersi in dettagli, in sottoproblemi di sottoproblemi. I veri problemi su cui si lavora — e che vengono sostenuti con tanta generosità — vengono mascherati e trasformati fino a quando le loro rappresentazioni si riducono a storielle, a innocue, innocenti, deliziose favolette.

Un esempio: uno studente di dottorato ha descritto il suo progetto di tesi come segue. Un bambino di sei o sette anni, sta seduto di fronte allo schermo di un computer su cui si vedono un gattino e un orsetto; il tutto a colori, naturalmente. Il gattino sta giocando con una palla. Il bambino dice al sistema computerizzato: «L'orsetto deve dire 'grazie' quando qualcuno gli dà qualcosa». Il sistema risponde con una voce sintetizzata, ma ciò nonostante piacevole a sentirsi: «Grazie, ho capito». Allora il bambino dice: «Micio, dai la tua palla al tuo amico». Immediatamente, vediamo il gattino sullo schermo lanciare la palla all'orsetto. Poi sentiamo l'orsetto dire: «Grazie, caro gattino». Questo è il nucleo di ciò che il sistema, il cui sviluppo costituirà il lavoro di tesi dello studente, deve saper fare. Da un punto di vista tecnico, il sistema deve capire istruzioni verbali — il che già di per sé non è un compito facile — e tradurle in un programma per il computer, che va poi integrato in modo lineare nella sua propria struttura di calcolo. Per niente banale, e, per di più, piuttosto toccante.

E ora una traduzione nella realtà. Un pilota di aereo da caccia viene interpellato dal suo sistema di guida automatica: «Signore, vedo una colonna di carri armati nemici là sotto. I Suoi ordini, prego». Il pilota: «Quando vedi qualcosa di quel tipo, non starmi a scocciare, distruggi quei bastardi e registra l'azione. È tutto». Il sistema risponde: «Sissignore!» e i missili dell'aereo partono verso terra.

Questo sistema di pilota automatico è uno dei tre sistemi d'arma che vengono descritti esplicitamente, per lo più come problemi di pertinenza dell'intelligenza artificiale, nella *Strategic Computing Initiative* («iniziativa di calcolo strategico»), uno dei nuovi grossi programmi di ricerca e sviluppo dei militari americani. Più di seicento milioni di dollari verranno spesi per questo programma nei prossimi quattro anni.

Non è mia intenzione, a questo punto, attaccare o oltraggiare il sistema militare. Voglio usare questo esempio tratto dalle abitudini della ricerca accademica sull'intelligenza artificiale negli Stati Uniti per illustrare gli eufemismi della dissimulazione linguistica, il cui effetto è quello di impedire il pensiero e, in ultima analisi, di mettere a tacere la coscienza.

Io non so proprio se sia in particolare la scienza dei computer, o la sua sottodisciplina «Intelligenza Artificiale», ad avere questa enorme affezione per l'eufemismo. Parliamo con una facilità così spettacolare di sistemi computerizzati capaci di capire, di vedere, decidere, esprimere giudizi, e così via, senza riconoscere la nostra stessa superficialità e incommensurabile ingenuità rispetto a questi concetti. Anestetizziamo la nostra capacità di valutare la qualità del nostro lavoro e, cosa più importante, di identificare e di prendere coscienza del suo uso finale.

Lo studente che ho citato immagina che il suo lavoro riguardi videogiochi per bambini, a base magari di gattini-giocattolo, orsacchiotti e palle. Il suo uso reale e premeditato probabilmente implicherà invece che un giorno o l'altro un giovane uomo, simile allo studente stesso e che ha dei genitori e forse una ragazza, verrà ucciso dall'esplosione di un missile lanciato contro di lui da un pilota automatico creato grazie alla ricerca dello studente. La distanza psicologica tra la concezione che lo studente ha del suo lavoro e le sue reali implicazioni è astronomica. È proprio questa enorme distanza che rende possibile non sapere e non chiedere se si stia facendo del lavoro sensato o se si stia contribuendo a una maggiore efficienza di apparecchiature criminali.

Non si può sfuggire a questa situazione senza domandarsi in continuazione: «Che cosa sto facendo in realtà? Quali sono l'uso e le applicazioni finali dei prodotti del mio lavoro?» e infine: «Sono contento o mi vergogno di aver contribuito a questo uso?».

In questo contesto mi torna alla mente un noto giornalista americano che, durante un dirottamento in Medio Oriente, suggerì che, se si fosse arrivati a una data situazione, gli israeliani uccidessero dieci prigionieri arabi tra quelli che detenevano in quel momento, e che se la situazione non fosse cambiata, ne uccidessero altri dieci il giorno dopo, e così via. Egli non avrebbe dovuto fare questa proposta, a meno di essere pronto ad andare personalmente tra i prigionieri, a guardare negli occhi gli uomini cui avrebbe poi detto: «tu e tu morirete oggi», e quindi a puntare la pistola alla testa di quelli scelti per essere uccisi, e a schiacciare il grilletto con il suo stesso dito.

Le stesse domande ce le dovremmo porre a proposito del nostro lavoro. Una volta abbandonati gli abbellimenti del nostro linguaggio, possiamo incominciare a parlare tra noi realisticamente e in modo serio del nostro lavoro di professionisti del computer.

> Tu, mio collega di molti anni di lavoro, stai lavorando su una macchina che è composta da dieci elevato alla quindicesima potenza, e più, microprocessori che lavorano contemporaneamente. Con l'aiuto di una macchina simile è possibile prima simulare e poi costruire bombe all'idrogeno molto più efficienti, più piccole e più leggere. Immagina, per un momento, di essere stato testimone oculare

a Hiroshima nel 1945, di aver visto morire persone cui era stata strappata la pelle. Vorresti che questo succedesse migliaia di altre volte, vorresti torturare in questo modo anche un solo essere umano con le tue mani? Se la risposta è no, indipendentemente da quale sarebbe il fine, allora devi smettere di fare questo lavoro.

Domande simili andrebbero fatte in riferimento ad altri rami della scienza dei computer, per esempio ai tentativi di rendere possibile ai computer la visione. Il progresso in questo settore verrà usato per guidare missili come il *Cruise* e il *Pershing* con sempre maggiore precisione verso i loro bersagli. E una volta giunti al bersaglio, sarà l'assassinio.

Ad affermazioni di questo tipo, si oppone spesso la dichiarazione che il computer è soltanto uno strumento. Come tale, può essere usato per il bene e per il male. Di per se stesso, non ha valori. Scienziati e tecnici non possono sapere come verranno applicati i prodotti del loro lavoro, se essi troveranno un'utilizzazione buona o malvagia. Quindi, scienziati e tecnici non possono essere ritenuti responsabili dell'applicazione finale del loro lavoro.

Quest'argomentazione la vedo manifestarsi nell'edificio vicino a quello in cui lavoro, il famosissimo Draper Laboratory. Questa istituzione si dedica quasi esclusivamente a problemi di guida dei missili e di navigazione dei sottomarini (un tempo, tra l'altro, faceva parte del Massachusetts Institute of Technology, MIT, l'università in cui insegno). Molti degli scienziati che vi lavorano fanno propria la dichiarazione appena citata. Dicono che i sistemi su cui lavorano possono portare uomini sulla Luna, e riportarli indietro, così come garantire che i missili puntati su Mosca colpiscano davvero Mosca, una volta lanciati. Non possono sapere in anticipo, dicono, per quale di questi o altri scopi il loro lavoro alla fine verrà usato. Come possono dunque venire considerati responsabili per le possibili conseguenze del loro lavoro? Lo stesso vale, nel complesso, per i professionisti del computer. Lo studente di dottorato che ho citato, che vuole riuscire a dialogare con lo schermo del suo computer, in realtà crede davvero che le applicazioni future del suo lavoro consisteranno esclusivamente in applicazioni innocenti come, per esempio, giocattoli per bambini. Magari la sua ricerca non è finanziata dalla *Strategic Computing Initiative* del Pentagono, magari non ha neppure mai sentito nominare l'*SCI*. Come gli si può quindi attribuire responsabilità per l'uso anti-umano che potrebbe venir fatto dei suoi risultati?

Qui arriviamo all'essenza della questione: oggi sappiamo praticamente con certezza che ogni risultato scientifico e tecnico verrà, se appena è possibile, utilizzato in sistemi militari. Il computer, insieme alla storia del suo sviluppo, è forse l'esempio-chiave. Ma a questo proposito dovremo anche pensare a tutto ciò che ha a che fare con il volo, con l'atomo, con i sistemi di comunicazione, con i satelliti, con le navicelle spaziali, e in generale alla maggior parte delle conquiste scientifiche del genio umano. Allora ci possiamo convincere del fatto che, nel mondo concreto in cui viviamo, l'onere della prova è su chi sostiene che un particolare nuovo sviluppo è immune dall'avidità dei militari.

In queste circostanze, i lavoratori della scienza e della tecnica non possono sfuggire alla responsabilità di informarsi sull'uso finale del loro lavoro. De-

vono quindi decidere, una volta saputo a che scopo verrà usato, se sarebbero disposti a perseguire quei fini con le loro stesse mani.

Penso che sia importante chiarire che io non credo che le istituzioni militari, di per se stesse, siano un male. Né affermerei che il fatto che una particolare tecnologia sia stata adottata dai militari sia, per questa sola ragione, un male. Nello stadio attuale dell'evoluzione dello stato-nazione sovrano (in altre parole, nel manicomio in cui viviamo) ogni Stato ha bisogno di un apparato militare, proprio come ogni città ha bisogno dei vigili del fuoco. Ma nessuno sostiene la necessità di una centrale dei pompieri a ogni angolo, e nessuno desidera avere vigili del fuoco che come attività collaterale vadano ad appiccare incendi nei paesi vicini alla città.

Ma noi assistiamo a una militarizzazione ogni giorno maggiore e più profonda di tutto il nostro mondo, e in particolare delle sue università e delle sue strutture scientifiche e tecnologiche. «Piccole» guerre scoppiano in quasi tutte le parti della Terra (in parte esse servono a provare le armi ad alta tecnologia delle «nazioni più avanzate»). Più della metà di tutti gli scienziati e gli ingegneri della Terra lavorano, più o meno direttamente, in istituzioni militari o in istituzioni sostenute principalmente dai militari. Questo è un male. A questo ci si deve opporre.

Dobbiamo anche riconoscere che è soltanto la nostra abitudine, ormai profondamente interiorizzata, di abbellire il nostro linguaggio per non urtare la nostra coscienza, che ci permette di parlare in termini di «armi» e di «vettori d'arma», mentre in realtà stiamo discutendo di esplosivi atomici e di proiettili contenenti bombe all'idrogeno. Queste non sono armi! Sono strumenti di assassinio di massa, ed è così che dovremmo chiamarli, in modo chiaro, distinto e senza circonlocuzioni. Quando ci si rende conto che uno strumento nucleare per l'assassinio di massa non è altro che una Auschwitz istantanea, un campo di sterminio istantaneo, una Auschwitz senza vagoni ferroviari o gente come Eichmann o il dottor Mengele, ma pur sempre una Auschwitz, è allora possibile lavorare su sistemi che guidano apparecchi di questo tipo verso città piene di vita? Questo, miei cari colleghi, è ciò che vi domando. Ma voi dovete porre seriamente a voi stessi domande di questo tipo, ed esaminare in profondità qualsiasi risposta troviate dentro di voi. Le vostre risposte devono infine manifestarsi nelle vostre azioni, concretamente in quello che fate ogni giorno.

Probabilmente, la malattia mentale più diffusa del nostro tempo è la credenza, condivisa in modo quasi universale, che l'individuo sia impotente. Questa illusione (che trova conferma in se stessa) verrà sicuramente indicata come contro-argomento per la mia tesi. Io chiedo, nevvero, che un'intera professione si rifiuti di partecipare alla follia criminale del nostro tempo. «Non può essere efficace» sento già che si dirà «Sì, se davvero nessuno lavorasse su quelle cose... ma è chiaramente impossibile. Dopo tutto, se non lo faccio io lo farà qualcun altro».

Per prima cosa, e al livello più elementare, devo dire che la regola «Se non lo faccio io, lo farà qualcun altro» non può servire come base di un comportamento morale. Qualunque delitto immaginabile può venire giustificato

sulla sua base. Per esempio: «se non rubo io i soldi all'ubriaco addormentato, lo farà qualcun altro».

Ma non è affatto banale stabilire il significato di «efficacia» in questo contesto. Di sicuro, l'efficacia non è una faccenda binaria, un *aut-aut*. È certo che, se quello che sto dicendo qui dovesse portare a uno sciopero da parte di tutti gli scienziati per quanto riguarda il lavoro sulle armi, questo andrebbe considerato un segno di efficacia. Ma ci sono molti altri livelli di efficacia, ben più modesti, cui io punto.

Penso che sia stato George Orwell a scrivere una volta: «Il massimo dovere degli intellettuali, di questi tempi, è affermare le verità più semplici nelle parole più semplici possibili» (tutti i tempi sembrano essere — è triste dirlo — «questi tempi», per quanto riguarda la saggezza di questa affermazione). Per me, ciò significa in primo luogo affermare l'assurdità del nostro mondo nelle mie azioni, nei miei scritti e con la mia voce. Spero in questo modo di stimolare i miei studenti, i miei colleghi, tutti coloro cui posso parlare direttamente. Spero in questo modo di incoraggiare coloro che hanno già cominciato a parlare in modo simile, e di essere incoraggiato da loro, e magari di risvegliare dal loro torpore tutti gli altri che riesco a raggiungere. Il coraggio, come la paura, è contagioso! Ogni successo, anche il più modesto, in questi tentativi deve anch'esso contare come esempio di efficacia. Inoltre, con i miei discorsi io inserisco ciò che discuto nell'agenda pubblica, e contribuisco alla sua legittimazione. Questi sono obiettivi modesti, che possono certamente essere raggiunti.

Ma, infine, voglio citare obiettivi più ampi come, per esempio:

> Liberare il mondo dagli ordigni nucleari di sterminio di massa, e forse anche dei generatori nucleari di potenza.
>
> Riorganizzare il mondo in modo che diventi impossibile convincere di nuovo i lavoratori di un Paese che sia per loro una necessità della vita quella di nutrire le loro famiglie della carne, del sangue e delle lacrime della gente di altri Paesi (è questo, purtroppo, il destino di molti lavoratori oggi; e non soltanto di quelli che si guadagnano il pane quotidiano nelle fabbriche di armamenti, ma altrettanto di quelli fra noi il cui lavoro quotidiano consiste nell'affinare armi di alta tecnologia).
>
> Riorganizzare il mondo in modo tale che ogni essere umano abbia a disposizione tutti i beni materiali necessari per vivere dignitosamente (ho sentito spesso persone ben intenzionate dire che, se applicheremo la tecnologia, e specialmente la tecnologia dei computer e delle comunicazioni, in modo saggio, potremo raggiungere questo obiettivo in qualcosa come cinquanta o cento anni. Ma lo possiamo raggiungere prima, e senza dover attendere progressi tecnologici. Perché l'ostacolo non è l'assenza di tecnologia, ma l'assenza di volontà politica!).

Una volta, ho sentito Elie Wiesel dire: «Dobbiamo credere che l'impossibile sia possibile». Io interpreto questo in due modi diversi.

Primo, se avessimo potuto credere che la «terra dei poeti e dei pensatori» potesse dare vita a fabbriche per lo sterminio umano, di efficienza paragonabile a quella delle fabbriche di automobili di Detroit, forse non avremmo dovuto assistere a Bergen Belsen. L'orrore impossibile si è dimostrato possibile ed è divenuto realtà.

Secondo, era «impossibile» nell'America di soli 150 anni fa abolire la schiavitù dei negri. Dopo tutto, l'intera economia del Sud degli Stati Uniti era basata sul cotone. Il cotone non poteva venir piantato né raccolto senza il duro lavoro, non pagato, di migliaia di esseri umani, dalla cui miseria il padrone delle piantagioni poteva ricavare il suo profitto. Ciò nonostante, all'inizio soltanto pochi uomini e donne lungimiranti — tutti sognatori — nel Massachusetts, e più tardi molti più cittadini, tra cui dei realisti, credettero che l'impossibile fosse possibile, che gli schiavi potessero essere liberati e la schiavitù eliminata. E diventò possibile. E diventò realtà.

Gli obiettivi impossibili che ho citato qui sono possibili, proprio come è possibile che distruggiamo la razza umana. Io da solo non posso né raggiungere il primo scopo, né impedire il verificarsi della seconda eventualità. Ma entrambi non possono realizzarsi senza di me. Non possono realizzarsi senza di noi.

Non ho il diritto di chiedervi nulla. Ma dal momento che voi, insieme ai vostri colleghi lavoratori nel mondo esterno, avete il potere di aumentare l'efficienza degli strumenti di sterminio di massa che abbiamo, e in questo modo di rendere più probabile l'assassinio dei miei figli; oppure di fermare per sempre l'attuale follia e, in tal modo, di aumentare la speranza, vostra come dei miei figli, di vivere in dignità umana, vi prego tutti: pensate a quello che realmente determinate col vostro lavoro, pensate a quali fini e in quale contesto ciò che fate verrà usato, chiedetevi se siete al servizio della vita o della morte.

INDICE DEI NOMI

ABM 45
Acton H. 116
ARPA 241
Albee E. 33
Andreski S. 151
Arendt H. 6, 29, 31
Aristotele 39
Armer P. 12
B-52 226
Bach J.S. 10, 38 s.
Bacone F. 31
Baseball, programma 176
Bateson G. 10
Bergler E. 121 s., 124
Bobrow 175
Boulding 29
Braun W. von 16
Bruner J. 199 s., 248
Bukharin N. 21, 26
Cavendish H. 32 n.
Cechov A.P. 33
Cheatham T. 12
Chomsky N. 6, 12, 132, 133 e n., 172, 222
Church A. 72, 75, 149
Clark B. 229
Colby K.M. 24 s., 168 s., 239 s.
Comfort 29
Crick F. 147 s.
Crocker J. jr 248 n.
Cruise 255
Darwin C.R. 38
Dendral, programma 208-210
Dennett D.C. 12
Dertounzos M. 6
Dewey J. 42, 46
Doctor, programma 24 s.
Dostoevskij F.M. 120 s.
Dreyfus H. 30
Eichmann A. 256
Einstein A. 147-149, 188
Eliot G.S. 77
Eliza, programma 7, 8, 22, 24, 26, 28 s., 169, 176-178, 185, 210
Ellul 29
Erikson E. 194 s.
Escher M.C. 10, 173
Evans M.A. 77
Everett C. 32 n.
Feigenbaum E.S. 152-154, 156, 165
Fein L. 8, 167
Feyerabend P.K. 7
Forrester J.W. 44 s., 221-225, 239
Freund P. 206 n.
Fromm E. 217
Galilei G. 39
Gandhi M.K. 222
Gödel K. 202 e n.
Green B.F. 175
Halle 222
Hao Wang 157
Heinsenberg W. 202
Henson 155
Hersch S. 216 n.
Hertz H. 115
Hitler A. 43 n., 217
Hofstaeder D. 10
Holmes O.W. 206
Horkheimer M. 225 s., 229
Hubbard R.L. 33
Hume D. 222
Huxley A. 125 s., 128
IBM 47, 82
Ibsen H. 33

Ionesco E. 186, 233
Jakobson R. 222
James W. 135
Johnson 33
Jonas H. 234 n.
Kaplan A. 135 s., 138
Kelly A. 13
Kelvin W.T. 211 s.
Kissinger H. 217
Langley 155
Lederberg J. 210
Leibniz G.W. 202
Lindsay R.K. 174
Lutero M. 36
Macsyma, programma 208-210
MARV 45
MIRV 45
Marcus S. 12, 136 n., 137 n.
Marshack A. 38, 40
Maslow A. 105 s.
Masters E.L. 33
McCarthy J. 186, 191, 234
McNamara, linea 243 s.
Mengele 256
Miller A. 33
Miller G.A. 150 s., 184, 189
Minsky M. 72, 149, 150 n., 211-213, 225
Mitscherlich A. 12
Moorer T. 216 e n., 217, 222
Morrison P. 214 s., 226
Moses J. 6, 210
Mozart W.A. 149
Mumford L. 6, 12, 29, 39 s.
Neumann J. von 143, 236
Newcomen 47
Newell A. 9, 133 s., 139, 156-159, 160 n., 161, 163-168, 182 s., 210, 224, 226, 242
Newton I. 183
Omero 38
Oppenheimer R. 253
Orwell G. 215, 257
Pearson K. 41
Pershing 255
Picasso P. 149
Platone 222
Platt J. 12
Poincarè H. 198 s.
Polanyi M. 21, 25 s., 79, 123 s., 133
Polya G. 158 s.
Priestely J. 32 n.
Putnam H. 12
Rayleigh J.W.S. 155
Reagan R. 252
Redlich F. 12
Reik T. 198
Ricardo D. 238
Richards I.A. 147
Roberts M.J. 234 s.
Roszach 29
Rubik, cubo di 18
SAGE 45
Student, programma 175 s.
Savery T. 47
Schank R.C. 9, 178-180, 182-185, 187, 192, 196
Scheffer I. 12
Selfridge O. 57 n.
Shakespeare W. 33
Shannon C. 190, 192
Shaw G.B. 176 n.
Shaw J.C. 156 s.
Simon H.A. 9, 126 s., 133 s., 139, 156-159, 160 n., 161, 163-169, 182 s., 187, 195, 210, 220 e n., 221 n., 223 s., 226, 228, 232, 240, 242
Skinner B.F. 32, 221, 223 s., 229, 239
Smith W.C. 247 n.
Spinoza B. 222
Stringfellow 155
Suppes P. 145
Terkel S. 231 s.
Trevithik 47
Turing A. 69, 72, 75, 105, 149
Turing, macchina di 7, 70-73, 75, 78, 81, 88, 90, 101, 108, 130, 160, 217
UNIVAC I 47 s.
Watson J.D. 147 s.
Weinberg S. 238 s.
Weiner H. 12
Whitehead A.N. 203
Wiener N. 210 s., 213, 239
Wiesel E. 257
Winner L. 234 n.
Winograd T. 9, 26, 180-182, 184, 192, 210, 232
Wright, fratelli 155
Wyck H. von 40

INDICE

Introduzione all'edizione italiana di Franco La Cecla 5
Prefazione alla prima edizione 11
Prefazione all'edizione inglese del 1984 15
Introduzione 21

1. Sugli strumenti 35
2. Da dove viene il potere del computer 53
3. Come funzionano i computer 81
4. La scienza e il programmatore coatto 113
5. Teorie e modelli 129
6. I modelli al computer in psicologia 147
7. Il computer e il linguaggio naturale 171
8. L'intelligenza artificiale 187
9. Programmi incomprensibili 207
10. Contro l'imperialismo della ragione strumentale 231

Appendice: Non senza di noi 251
Indice dei nomi 259